

Entwicklung eines Bildverarbeitungssystems zur Überwachung des Locomotion Scoring in Milchviehbetrieben

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum agriculturarum

(Dr. rer. agr.)

im Fach Agrarwissenschaften

eingereicht an der

Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

M.Sc. Sarah Dieckmann

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

Prof. Dr. Bernhard Grimm

Gutachter/innen

1. PD Dr. rer. agr. habil. Anke Römer

2. Prof. Dr. vet. med. Jörn Karge

3. Prof. Dr. agr. habil. Norbert Kanswohl

Tag der mündlichen Prüfung: 25.03.2019

Meinem Opa

I. Danksagung

Zuerst möchte ich mich beim Bundesministerium für Wirtschaft und Energie bedanken, die dieses Forschungsprojekt („LoBoScore“, Reg.-Nr. 16KN016247) über das Zentrale Innovationsprogramm Mittelstand gefördert haben. Im selben Zuge bedanke ich mich auch bei allen Partnern und Kollegen, ohne welche die Umsetzung des Projekts nicht möglich gewesen wäre. Ein besonders großer Dank gilt an dieser Stelle dem Team der Gesellschaft für angewandte Informatik e.V., die für die Softwareentwicklung in dem Projekt verantwortlich waren - insbesondere Duy Tu Diep für die konstruktive Zusammenarbeit. Ebenso danke ich Lars Schmidt, in dessen Stall das System installiert wurde und der alle Baumaßnahmen und Versuche geduldig begleitet hat, für seine freundliche Kooperation.

Bei Dr. agr. M. Krockert vom Albrecht Daniel Thear-Institut für Agrar- und Gartenbauwissenschaften und Prof. S. Heinz von der Frankenförder Forschungsgesellschaft mbH bedanke ich mich für die Ermutigung eine Dissertation zu diesem Thema zu verfassen.

Frau PD Dr. agr. habil A. Römer von der Landesforschungsanstalt für Landwirtschaft und Fischerei MV möchte ich für die gute Betreuung der Arbeit, die Ermutigungen in den richtigen Momenten und das allzeit offene Ohr bei Problemen danken.

Herrn Dr. nat. Wolfgang Waser von der Frankenförder Forschungsgesellschaft mbH möchte ich für die Unterstützung bei der R-Programmierung und statistischen Analyse danken.

Ein großer Dank gilt meinen Eltern, ohne deren langjährige Unterstützung die Promotion nicht möglich gewesen wäre.

Andreas und Annika danke ich für das Korrekturlesen und Tim für den Glauben an mich und die nicht minderwichtigen Aufmunterungen und das allzeit offene Ohr.

II. Zusammenfassung

Ein System zur automatischen Erfassung der Klauengesundheit von Milchkühen kann durch möglichst frühzeitige präventive und kurative Eingriffe sowohl das Tierwohl steigern als auch ökonomische Verluste minimieren. Ziel war die Entwicklung eines automatischen Systems, das mittels 3D Sensordaten die Rückenkrümmung einer Kuh erfassen kann und einen entsprechenden Locomotion Score ausgibt.

In dieser Studie wurden von August bis Dezember 2016 auf einem Milchviehbetrieb mit Boxenlaufstallhaltung 200 Kühe der Rasse Deutsche Holstein automatisch von einem System nach Verlassen des Melkstandes erfasst. Das System bestand aus zwei RFID-Antennen zur Tiererkennung, einem 3D Sensor zur Lahmheitsbeurteilung und zwei 2D Kameras zur visuellen Bonitur und Validierung. Anhand der 2D Videos wurde nach der Methode von SPRECHER et al. (1997) für 1489 Datensätze ein Locomotion Score vergeben.

Im 2,5D-Tiefenbild wurden zunächst die Hüfthöcker als höchste, äußerste Punkte im Bild identifiziert. Ausgehend von einer Verbindungslinie zwischen den Hüfthöckern, wurde die Wirbelsäule als Polylinie der höchsten Punkte in Richtung Bildanfang ermittelt. Im 3D-Raum wurde dann an die gefundene Rückenlinie ein Kreis approximiert, dessen Radius umgekehrt proportional zur Krümmung des Kuhrückens ist.

Der Vergleich der visuell vergebenen und der automatisch gemessenen Scores ergab für das Modell mit fünf Stufen eine Übereinstimmung von 59 % und für das Modell mit drei Stufen 65 %. Die Korrelation nach Spearman ($r_s = 0,73$) zeigte einen deutlichen Zusammenhang zwischen den berechneten und den visuellen Scores. Mit einer Genauigkeit von 82 % wurden 96 % der lahmen Kühe auch als lahm erkannt. Die größte Übereinstimmung wurde für den Locomotion Score 2 erreicht (80 %).

In dieser Arbeit wurde ein automatisches System entwickelt, welches durch die Berechnung der Krümmung der Rückenlinie insbesondere beginnende Lahmheiten sehr gut erkennen kann.

III. Abstract

A system for automatic detection of lameness in dairy cows can increase animal welfare as well as minimize economic losses through early preventative and curative treatments. The aim of this study was to develop an automatic system that uses 3D sensor data to detect the spine curve of a cow and generate a corresponding locomotion score.

In the present study 200 German Holstein cows housed in a free stall barn were recorded via an automatic system after leaving the milking parlour from August to December 2016 on a dairy farm. The system consisted of two RFID-antennas, a 3D depth camera for automatic lameness detection and two standard 2D cameras for visual classification and validation. An amount of 1,489 2D sets of data were locomotion scored according to the method of SPRECHER et al. (1997). Locomotion Score 1 was used to score a healthy cow and Locomotion Score 2 – 5 for lameness with increasing severity. In the analysis five levels (score 1 – 5) as well as three levels (score 1, 2, and ≥ 3) were compared.

In the 2.5D depth image the hip joints were identified as the highest, outermost points in the image. Starting from a connection line between the hip joints, the spine was determined as a polyline of the highest points in anterior direction. On the 3D surface, a circle was approximated to the detected back line whose radius is inversely proportional to the curvature of the cow's back.

A comparison of the visually assigned and automatically detected scores showed a 59 % match of the five-level model and 65 % for the three-level model. The Spearman correlation ($r_s = 0.73$) showed a clear correlation between the calculated and the visual scores. With an accuracy of 82 %, 96 % of lame cows were identified as lame. The highest agreement was reached for locomotion Score 2 (80 %).

In the present study an automatic system was developed, which can recognize in particular beginning lameness by calculating the curvature of the back.

IV. Inhaltsverzeichnis

I.	Danksagung	i
II.	Zusammenfassung	ii
III.	Abstract.....	iii
V.	Abbildungsverzeichnis	vii
VI.	Tabellenverzeichnis	ix
VII.	Abkürzungen	xi
1.	Einleitung und Zielstellung	- 1 -
2.	Literatur	- 3 -
2.1	Lahmheit.....	- 3 -
2.2	Normale Bewegungsmuster von Kühen und die häufigsten Abweichungen	- 4 -
2.3	Locomotion Scoring	- 6 -
2.3.1	Durchführung des visuellen Locomotion Scorings	- 6 -
2.3.2	Lahmheitsindikatoren des visuellen Locomotion Scorings.....	- 7 -
2.3.3	Subjektivität der Beobachter	- 10 -
2.4	Verhaltens- und Leistungsänderungen durch Lahmheit.....	- 10 -
2.5	Automatische Erfassung des Locomotion Scores	- 11 -
2.5.1	Kinetische und kinematische Ansätze	- 12 -
2.5.2	Lahmheitsbeurteilung anhand von Videoanalysen.....	- 14 -
2.5.3	Indirekte Messungen	- 16 -
2.5.4	Versuchsaufbau und Messtechnik zur automatischen Erfassung des Locomotion Scores.....	- 17 -
2.5.5	Validierung automatischer Locomotion Scores	- 28 -
3.	Material und Methode	- 32 -
3.1	Tiermaterial	- 32 -
3.2	Versuchsumgebung	- 32 -

Inhaltsverzeichnis

3.2.1	Haltung	- 33 -
3.2.2	Fütterung.....	- 34 -
3.2.3	Klauenmanagement	- 34 -
3.3	Technik	- 35 -
3.3.1	3D Sensor	- 35 -
3.3.2	Tiererkennung.....	- 36 -
3.3.3	2D Kameras	- 36 -
3.4	Visuelles Locomotion Scoring	- 37 -
3.5	Statistische Auswertung	- 37 -
4.	Ergebnisse.....	- 39 -
Prozessentwicklung		- 39 -
4.1	Entwicklung des experimentellen Designs.....	- 39 -
4.2	2,5D- und 3D-Datenverarbeitung	- 44 -
4.3	Eruierung von Messpunkten	- 45 -
4.3.1	Berechnung 1	- 45 -
4.3.2	Berechnung 2	- 46 -
4.4	Berechnung der automatischen Scores	- 47 -
4.5	Visuelles Locomotion Scoring	- 48 -
Validierung und Klassifizierung der Algorithmen zur Ableitung des Locomotion Scores.....		- 49 -
4.6	Automatisches Locomotion Scoring	- 49 -
4.6.1	Auswertung 1	- 49 -
4.6.2	Auswertung 2	- 50 -
4.6.3	Auswertung 3	- 53 -
4.6.3.1	Stichprobe 3	- 53 -
4.6.3.2	Stichprobe 4	- 55 -
4.6.3.3	Stichprobe 5	- 57 -
4.6.3.4	Stichprobe 6	- 61 -

Inhaltsverzeichnis

5.	Diskussion	- 64 -
	Prozessentwicklung	- 64 -
5.1	Experimentelles Design	- 64 -
5.2	Datenverarbeitung	- 68 -
5.3	Eruierung von Messpunkten	- 69 -
5.4	Visuelles Locomotion Scoring	- 71 -
	Validierung und Klassifizierung der automatischen Scores	- 74 -
5.5	Auswertung 1	- 75 -
5.6	Auswertung 2	- 75 -
5.7	Auswertung 3	- 77 -
5.7.1	Stichprobe 3	- 77 -
5.7.2	Stichprobe 4	- 78 -
5.7.3	Stichprobe 5	- 80 -
5.7.4	Stichprobe 6	- 82 -
6.	Schlussfolgerungen	- 84 -
7.	Literaturverzeichnis	- 86 -
8.	Anhang	- 94 -
9.	Selbstständigkeitserklärung	- 95 -

V. Abbildungsverzeichnis

Abbildung 1: Analyse der beim Gehen entstehenden Abdrücke; Bezeichnung der Klauen: FL=vorne links, FR=vorne rechts, RL=hinten links, RR=hinten rechts (aus: TELEZHENKO UND BERGSTEN 2005).....	- 4 -
Abbildung 2: Flussdiagramm der Methode von POURSAHERI et al. (2010). Input des Systems sind Videoaufzeichnungen und es wird ein numerischer Wert ausgegeben, der im Verhältnis zur Rückenkrümmung steht. Dafür wird die Kuh vom Hintergrund separiert und anhand des Binärbildes die Rückenkrümmung berechnet.....	- 21 -
Abbildung 3: Parameter α_1 , α_2 , α_3 und L_1 einer nachgebildeten Rückenlinie einer Kuh mittels 3D Sensor (VIAZZI et al. 2014).....	- 24 -
Abbildung 4: Darstellung der automatischen Berechnung der Wirbelsäule und der Hüfthöcker aus einem 3D Bild einer Kuh (JABBAR et al. 2017). Das erste Bild zeigt ein unbearbeitetes Tiefenbild, gefolgt von demselben Bild ohne Hintergrund. Darauf folgt die Krümmungsberechnung mit dargestellten Höhepunkten, welche anschließend in ein Binärbild umgewandelt werden und so die Wirbelsäule sichtbar machen. Am Ende steht ein Auswahlbild auf welchem die Krümmungen (höchste konvexe Regionen d.h. Wirbelsäule und Hüfthöcker) deutlich sichtbar sind.....	- 25 -
Abbildung 5: Krümmungsdaten berechnet aus einem Tiefenbild einer Kuh (HANSEN et al. 2018). Vor (links) und nach (rechts) der automatischen Ausrichtung des Bildes. Die am stärksten gekrümmten Bereiche werden deutlich gezeigt.	- 28 -
Abbildung 6: Skizze der Versuchsumgebung mit Darstellung der Laufwege der Kühe	- 33 -
Abbildung 7: Asus xTion Pro Live.....	- 35 -
Abbildung 8: Antennen der Tiererkennung im Treibgang	- 36 -
Abbildung 9: RFID Sensor am Halsband einer Kuh	- 36 -
Abbildung 10: Im Stall montierte 2D Kamera.....	- 37 -
Abbildung 11: Skizze der Versuchsumgebung in der ersten Versuchsphase. Darstellung der Position des 3D Sensors und der Beobachterin im ersten Versuchszyklus	- 39 -
Abbildung 12: Skizze der Versuchsumgebung in der zweiten Versuchsphase. Darstellung der Position der 3D Sensoren, 2D Kameras und der Tiererkennung	- 40 -

Abbildungsverzeichnis

Abbildung 13: Schutzgehäuse des 3D Sensors - Aufnahme oberhalb der Decke am Versuchsende	- 41 -
Abbildung 14: Schutzgehäuse des 3D Sensors von unten durch die Decke aufgenommen. Die Aussparung im Aluminiumgehäuse ist mit einer Isolier-Fensterglasscheibe abgedeckt. Darüber ist der 3D Sensor.....	- 41 -
Abbildung 15: Darstellung der Anordnung der Messtechnik im Schutzgehäuse	- 42 -
Abbildung 16: Darstellung der Ordnerstruktur für die Datenerfassung	- 43 -
Abbildung 17: 2,5D-Tiefenbild einer Kuh mit algorithmisch identifizierter Hüft- und Wirbelsäulenlinie in Rot (Ansicht von oben). Die Hüftlinie ist die Verbindungslinie der beiden höchsten, äußerten Punkte. Die Wirbelsäule wird als Polylinie der höchsten Punkte ausgehend von der Hüftlinie ermittelt.....	- 46 -
Abbildung 18: Berechnung des Endpunkts der Rückenlinie. Abstandmessung der Breite des Kuhhalses im Abstand von 2 cm. Weisen vier Messungen hintereinander positive Werte auf, wird diese Stelle als Endpunkt definiert.	- 46 -
Abbildung 19: 3D-Dreiecknetz einer Kuh mit der gefundenen Rückenlinie in Rot und der approximierten Rückenkrümmung in Gelb (Seitenansicht) in der 3D-Software „Final Surface“	- 47 -
Abbildung 20: Vergleich der Häufigkeiten der einzelnen Scores bei der Berechnung mit Hals und ohne Hals jeweils für die Stichprobe von 197 Datensätzen (Stichprobe 2)	- 51 -
Abbildung 21: Darstellung der Messwerte der Rückenkrümmung aus Berechnung 3 zu den einzelnen visuellen Scores aus Stichprobe 3 ...	- 54 -
Abbildung 22: Box-Plot der Messwerte der Rückenkrümmung aus Berechnung 3 zu den visuellen Scores anhand von Stichprobe 4	- 55 -
Abbildung 23: Box-Plot der Messwerte der Rückenkrümmung aus Berechnung 3 zu den visuellen Scores anhand von Stichprobe 5	- 58 -
Abbildung 24: Darstellung der Messwerte der Rückenkrümmung aus Berechnung 3 in den einzelnen Scores anhand von Stichprobe 6.....	- 62 -

VI. Tabellenverzeichnis

Tabelle 1: Locomotion Scoring nach SPRECHER et al. (1997).....	- 8 -
Tabelle 2: Lahmheitsbeurteilungssystem nach THOMSEN et al. (2008).....	- 9 -
Tabelle 3: Erste Klassifizierung der Krümmungsgrade.....	- 48 -
Tabelle 4: Übersicht über die für die Berechnungen und Auswertungen zugrundeliegenden Datensätze und Stichproben	- 49 -
Tabelle 5: Darstellung der Ergebnisse aus Berechnung 1 (Polylinie der Wirbelsäule ohne definierten Endpunkt) anhand von Stichprobe 1. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores	- 50 -
Tabelle 6: Darstellung der Ergebnisse aus Berechnung 2 (Polylinie der Wirbelsäule mit definiertem Endpunkt) anhand von Stichprobe 2. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores - „Auswertung ohne Hals“	- 50 -
Tabelle 7: Darstellung der Ergebnisse aus Berechnung 1 (Polylinie der Wirbelsäule ohne definierten Endpunkt) anhand von Stichprobe 2. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores - „Auswertung mit Hals“	- 51 -
Tabelle 8: Anzahl der Abweichungen der automatischen Scores vom visuellen Score je Größe der Abweichung und je visuellem Score anhand von Stichprobe 2 – „Auswertung mit Hals“ (Berechnung 1).....	- 52 -
Tabelle 9: Anzahl der Abweichungen der automatischen Scores vom visuellen Score je Größe der Abweichung und je visuellem Score anhand von Stichprobe 2 – „Auswertung ohne Hals“ (Berechnung 2)	- 52 -
Tabelle 10: Klassifizierung mit optimierten Grenzwerten anhand der Ergebnisse von Berechnung 2	- 53 -
Tabelle 11: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 3. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores	- 54 -
Tabelle 12: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 4. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores	- 55 -
Tabelle 13: Statistische Auswertung von Berechnung 3 anhand von Stichprobe 4 ...	- 56 -
Tabelle 14: Test auf Selektivität und Sensitivität von Berechnung 3 anhand von Stichprobe 4	- 57 -

Tabellenverzeichnis

Tabelle 15: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 5. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores	- 57 -
Tabelle 16: Statistische Auswertung von Berechnung 3 anhand von Stichprobe 5 ...	- 59 -
Tabelle 17: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 5. Anzahl der Übereinstimmungen und Abweichungen der visuellen und automatischen Scores innerhalb der visuellen Scores	- 60 -
Tabelle 18: Test auf Spezifität und Sensitivität von Berechnung 3 anhand von Stichprobe 5	- 60 -
Tabelle 19: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 6. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores	- 61 -
Tabelle 20: Statistische Auswertung von Berechnung 3 anhand von Stichprobe 6 ...	- 62 -
Tabelle 21: Test auf Spezifität und Sensitivität von Berechnung 3 anhand von Stichprobe 6	- 63 -

VII. Abkürzungen

aLS	automatisches Locomotion Scoring / visueller Locomotion Score
AUC	Area Under the Receiver Operating Characteristics Curve
diff	Differenz
DF	Anzahl der Freiheitsgrade
F value	F-Wert
LS	Locomotion Scoring / Locomotion Score
lwr	Untergrenze der Konfidenzintervalle
Mean Sq	mittlere Abweichungsquadrate
N	Anzahl
P adj	Signifikanz
Pr (>F)	Signifikanz
RFID	radio-frequency identification = Identifizierung mit Hilfe elektromagnetischer Wellen
Sum Sq	Quadratsumme
TMR	Totale Mischration
upr	Obergrenze der Konfidenzintervalle
vLS	visuelles Locomotion Scoring / visueller Locomotion Score

1. Einleitung und Zielstellung

Gesunde Klauen sind die Grundlage einer produktiven Nutzungsdauer, effektiven Fortpflanzung und hohen Milch- und Fleischleistung (GÜNTHER 1991). Die Klauen der Milchkühe richtig zu pflegen und gesund zu erhalten, ist demnach - nicht nur aus ökonomischer, sondern auch aus Sicht des Tierwohls - ausgesprochen wichtig. Finanzielle Verluste entstehen durch prophylaktische und therapeutische Maßnahmen, aber auch infolge von Mehrkosten bzw. Einbußen durch reduzierte Fruchtbarkeit, verringerte Milchleistung und vorzeitige Schlachtung. Denn lahme Kühe sind anfälliger für Stoffwechselerkrankungen und Euterentzündungen und zeigen häufig verschlechterte Fruchtbarkeitsparameter (SOGSTAD et al. 2006, BICALHO et al. 2007, GREEN et al. 2014). In Deutschland stehen Klauen- und Gliedmaßenkrankungen auf Platz drei im Ranking der häufigsten Abgangsursachen (ADR 2017). Im modernen Herdenmanagement sind daher die Klauenpflege und Lahmheitsbeurteilung sowie die sofortige Untersuchung und Behandlung als lahm identifizierter Tiere die wichtigsten Maßnahmen für die Erhaltung und Kontrolle der Klauengesundheit (SPRECHER et al. 1997, KOFLER 2014). Deshalb ist es besonders wichtig, ein automatisches System in die landwirtschaftlichen Betriebe zu integrieren, mit welchem Klauenerkrankungen frühzeitig erkannt werden, um präventive und kurative Eingriffe möglichst frühzeitig durchführen zu können. Dadurch könnten nicht nur die ökonomischen Verluste minimiert werden, sondern auch das Tierwohl gesteigert.

Bisher werden Lahmheiten in Praxisbetrieben - wenn überhaupt - mittels Pedometer automatisch überwacht, welche aber in erster Linie zur Brunstüberwachung eingesetzt werden. Die Entwicklung videobasierter Lahmheitserkennung ist noch in ihren Anfängen. VIAZZI et al. (2014), VAN HERTEM et al. (2014) und JABBAR et al. (2017) haben zwar bereits mittels 3D-Daten die Rückenkrümmung von Kühen erfasst, jedoch konnte sich bisher keine veröffentlichte Methode für den täglichen Gebrauch in Praxisbetrieben durchsetzen.

Ziel der Dissertation ist es, ein automatisches System zur Erkennung der Rückenlinie mittels 3D-Datenerfassung direkt in einem Praxisbetrieb zu entwickeln und zu testen, um Lahmheiten frühzeitig zu erkennen und Läsionen an Klauen und Gelenken abzuwenden. Die praktische Bearbeitung erfolgte in einem landwirtschaftlichen Betrieb in Brandenburg vom 01.06.2015 bis 31.12.2016. Das Ziel der Arbeit lässt sich

unterteilen in zwei Teilziele. Zum einen die Prozessentwicklung und zum anderen die Klassifizierung und Validierung der entwickelten Algorithmen.

Folgende Zielstellungen ergeben sich für die Dissertation:

1. Prozessentwicklung

- Entwurf der einzelnen Versuchsdesigns für die einzelnen Projektphasen sowie Vorbereitung und Durchführung der Versuche unter Praxisbedingungen
- Analyse von Prozessen und Arbeitsvorgängen zur Erhöhung der Effizienz und Funktionssicherheit in der Praxis sowie zur Erkennung und Beseitigung möglicher Schwachstellen und Fehlerquellen
- Eruierung der Messpunkte für das Locomotion Scoring

2. Validierung und Klassifizierung der Algorithmen zur Ableitung des Locomotion Scores

- Entwicklung und Festlegung von Referenzgrößen für die einzelnen Scores
- Untersuchung von Korrelationen zwischen Sensorsignalen und sicher differenzierbaren Gesundheitszuständen

2. Literatur

2.1 Lahmheit

Lahmheit kann definiert werden als eine klinische Ausprägung von schmerzverursachenden Erkrankungen, welche hauptsächlich den Bewegungsapparat betreffen und dadurch eine unnatürliche Bewegung beziehungsweise Abweichungen von einem normalen Gangbild und einer gesunden Haltung nach sich ziehen. Die Ausprägung der Lahmheit reicht von einer Steifheit oder Asymmetrie in der Beinbewegung bis zur völligen Schonung eines Beins oder sogar dem Unvermögen aufzustehen (STRAW et al. 1999, MAAS 2009).

Schmerzen in den Gliedmaßen lösen in den meisten Fällen eine Lahmheit aus, können aber auch an anderen Stellen des Körpers lokalisiert sein (PETERSE 1992). Bei mehr als 90 % der Lahmheiten ist diese in Läsionen der Klaue begründet (BERGSTEN 2001). Wird der Schmerz aus einem anderen Teil des Bewegungsapparates verursacht, betrifft das Problem meist nur das individuelle Tier, wohingegen Klauenprobleme die gesamte Herde betreffen können (PETERSE 1992).

Durch Lahmheiten können finanzielle Verluste direkt durch prophylaktische und therapeutische Tierarzt- und Materialkosten entstehen, aber auch indirekt durch Milchverluste, verlängerte Zwischenkalbezeiten, gesenkte Nutzungsdauer, reduzierte Schlachterlöse und nicht zuletzt aus dem erhöhten Arbeitsaufwand (MÜLLER 2011, POLITIEK et. al. 1986, VAN STRATEN et al. 2011). Da ein Großteil der Lahmheiten in der Frühlaktation auftritt, ist es nicht oder kaum vermeidbar, dass die Fruchtbarkeit und die Milchleistung nachteilig beeinflusst werden und sich so negativ auf die Ökonomie des Milchviehbetriebes auswirken (REHAGE et al. 2011). In den Untersuchungen von WIEDENHÖFT (2005) lagen die Kosten pro Tier und Trächtigkeit bei lahmen Kühen durch eine verlängerte Güstzeit im Schnitt um rund 49 Euro höher als bei nicht lahmen Kühen. Nach MÜLLER (2011) können sich diese Verluste bei lahmen Kühen in Einzelfällen auf über 600 Euro summieren.

TADICH et al. (2010) haben den Zusammenhang zwischen Locomotion Score und Klauenerkrankungen untersucht. Dabei wurde das visuelle Locomotion Scoring nach SPRECHER et al. (1997) durchgeführt und alle Kühe mit einem Score > 1 als lahm bezeichnet. Die Krankheiten, die einen besonders schlechten Locomotion Score (LS) nach sich ziehen, sind Klauensohlengeschwür und Doppelsohle. Da andere

Erkrankungen nicht zwingend einen schlechten Locomotion Score als Folge haben, haben TADICH et al. (2010) gemutmaß, dass das Locomotion Scoring vielleicht nicht sensibel genug ist, um alle Klauenerkrankungen aufzuzeigen.

2.2 Normale Bewegungsmuster von Kühen und die häufigsten Abweichungen
 Idealerweise zeigt eine Kuh eine gute Vorwärtsbewegung. Dabei ist die Schrittlänge der Größe des Tieres angemessen und bei allen vier Beinen gleich lang. Die Schrittlänge beschreibt die Strecke zwischen zwei aufeinanderfolgenden Abdrücken desselben Hinterfußes (Abbildung 1).

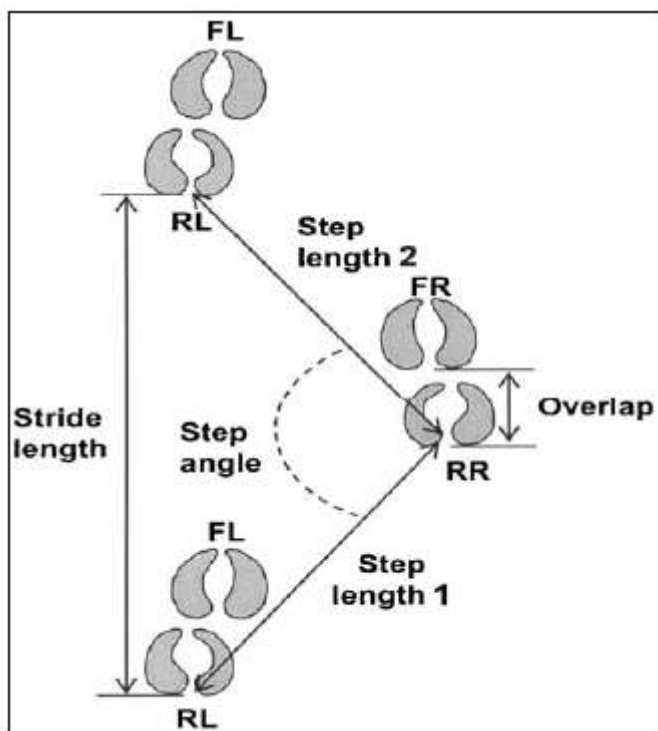


Abbildung 1: Analyse der beim Gehen entstehenden Abdrücke; Bezeichnung der Klauen: FL=vorne links, FR=vorne rechts, RL=hinten links, RR=hinten rechts (aus: TELEZHENKO UND BERGSTEN 2005)

Die Schrittweite, ergibt sich aus dem Abstand zwischen der rechten und der linken Hinter- bzw. Vorderklaue (TELEZHENKO 2005). Diese richtet sich im Normalfall nach dem Fundament des jeweiligen Tieres.

Das Übertreten, oft auch „overlap“ oder „tracking-up“ genannt, bezeichnet die Längsstrecke zwischen dem vorderen Fußabdruck und dem nächsten Abdruck des hinteren Fußes der gleichen Seite. Je weiter hinten die Hinterhand hinter der

Vorderhand auftritt, desto schlechter ist dieses zu bewerten (TELEZHENKO 2005, FLOWER UND WEARY 2006).

Der Schrittwinkel stellt den Winkel zwischen den Verbindungslinien dreier aufeinanderfolgender Abdrücke der Hinterfüße dar. Geht eine Kuh auf hartem Untergrund, deutet ein verkleinerter Winkel auf eine Lahmheit hin. Auf weichem Untergrund vergrößern lahme Kühe den Schrittwinkel hingegen (TELEZHENKO UND BERGSTEN 2005).

Bei einem normalen Gang handelt es sich um einen klaren Vier-Takt (FLOWER UND WEARY 2006). Ist dieser Rhythmus in Richtung Drei-Takt gestört, so liegt wahrscheinlich eine Lahmheit vor.

Die Kopfbewegung einer gesunden Kuh ist gleichmäßig und eben. Im gesunden Zustand sollte sich dieser während der Bewegung nur minimal heben und senken. Lahmt ein Tier auf einer Gliedmaße, so wird die Kopfbewegung unregelmäßig und bei Belastung der erkrankten Gliedmaße deutlicher sichtbar (FLOWER UND WEARY 2006).

Die Rückenlinie ist für die Gangbeschreibung und Beurteilung bei Milchkühen von großer Bedeutung. Bei einer gesunden Kuh ist der Rücken gerade, wohingegen eine lahme Kuh ihren Rücken aufwölbt (SPRECHER et al. 1997, FLOWER UND WEARY 2006).

Die Ab-/Adduktion ist die Abweichung der Bewegung des Fußes in der Vorwärtsbewegung von der Bewegungsrichtung (ESPEJO et al. 2006). Solche Abweichungen in Form von Rotationen können auf Entlastungen hindeuten. Außerdem können ständige Fehlstellungen Vorboten für spätere Lahmheiten sein. Durch abnormes Auftreten mit der Klaue kann es zu Fehlbelastungen kommen, wodurch Gewebe in der Klaue oder Gelenke in der Gliedmaße beschädigt werden können.

Auch die Beweglichkeit der Gelenke ist ein Aspekt, der den Gang einer Milchkuh beeinflusst. Im Idealfall lassen sich die Gelenke ausreichend beugen und strecken. Ist dies nicht der Fall, gehen die Tiere steif (FLOWER UND WEARY 2006). Diese Beweglichkeit bezieht sich besonders auf das Sprunggelenk, da hier Versteifungen am häufigsten vorzufinden sind. Hat eine Kuh eine geringe Beweglichkeit im Sprunggelenk, hat dies auch Auswirkungen auf das Übertreten, denn durch die Steife wird auch die Schrittlänge verkürzt.

2.3 Locomotion Scoring

Die Verbesserung des Tierwohls ist Bestandteil vieler Forschungsprojekte. Die Lahmheitserkennung und -vermeidung spielt hier oft eine wichtige Rolle auf dem Gebiet der Klauengesundheit, denn Lahmheiten sind ein offensichtliches Zeichen für Klauenerkrankungen oder andere schmerzhaftes Beinverletzungen, welche das Tierwohl beeinträchtigen. Das Locomotion Scoring wird als Instrument zur Erkennung von Klauenerkrankungen und Schmerzen in den Gliedmaßen eingesetzt (WELFARE QUALITY 2009, AHDB DAIRY 2017).

Beim Locomotion Scoring wird die Bewegung einer Kuh qualitativ beurteilt. Für Milchkühe wurden unterschiedliche Ansätze entwickelt und untersucht, um eine Lahmheit frühestmöglich zu erkennen. Bezüglich des Ganges werden besonders Veränderungen der Gliedmaßen betrachtet. Diese können sich in einem asymmetrischen Gangbild oder einer ungleichmäßigen Gewichtsverteilung auf den Klauen bzw. einzelnen Gliedmaßen zeigen, aber auch die Schrittlänge und die Beweglichkeit in den Gelenken werden beobachtet. Unter dem Kriterium Körperhaltung werden die Krümmung des Rückens und das Nicken des Kopfes während des Laufens betrachtet. Zudem können auch andere Faktoren, wie die Laufgeschwindigkeit oder Probleme bei Richtungsänderungen oder beim Aufstehen, analysiert werden. Zumeist wird zur Bewertung des Locomotion Scores eine Ordinalskala verwendet, die zwischen zwei und neun Stufen haben kann. Auch eine stufenlose Bewertung des Locomotion Scores ist möglich, wird aber selten durchgeführt, da diese Methode schwerer zu erlernen ist und sich in der Praxis nicht so leicht anwenden lässt (SCHLAGETER-TELLO et al. 2014).

In den letzten Jahrzehnten wurden verschiedene Ansätze entwickelt, um die Lahmheitsbewertung sowohl subjektiv durch einen Beobachter als auch objektiv durch computergestützte Analysemethoden zu bewerten. Im Folgenden werden ausgewählte Studien beispielhaft vorgestellt.

2.3.1 Durchführung des visuellen Locomotion Scorings

Die entwickelten Methoden berücksichtigen unterschiedliche Faktoren für die Beurteilung und unterscheiden sich in der Anzahl möglicher Stufen zur Bewertung. Alle Methoden sind rein subjektiv und beschreiben keine diagnostischen Maßnahmen. Die Bewertung erfolgt numerisch in Form von Noten. Die Gegebenheiten vor Ort für die Durchführung des Locomotion Scorings werden meist ähnlich beschrieben. Die Kuh

sollte auf einem geraden und ebenen Untergrund laufen, der nicht rutschig ist. Außerdem sollte sie in ihrem eigenen Tempo laufen und nicht getrieben werden. Der Beobachter betrachtet das Tier dabei aus einem gewissen Abstand von der Seite (MANSON UND LEAVER 1988, SPRECHER et al. 1997, WHAY et al. 1997, WINCKLER UND WILLEN 2001). Diese Beschreibung lässt sich zurückführen auf die einheitlichen Produktionsbedingungen und -abläufe in Milchviehbetrieben. Es bietet sich an, das Locomotion Scoring nach dem Melken durchzuführen, da die Kühe dann meist einzeln und ohne Zwang auf befestigtem Untergrund laufen.

2.3.2 Lahmheitsindikatoren des visuellen Locomotion Scorings

Die Gangbeurteilung von MANSON UND LEAVER (1988) gilt als Grundlage der meisten nachfolgend beschriebenen Gangbeurteilungssysteme in der Literatur. Sie geben eine Skala von 1 bis 5 vor, auf welcher auch die Vergabe von halben Noten möglich ist. Bei der Benotung der Kühe wird nicht nur Wert auf die Schrittverkürzung gelegt, sondern auch auf die Gliedmaßenführung in der Bewegung (Ab-/Adduktion) und die Gewichtsverteilung auf die Klauen. Zusätzlich werden das Gangbild in der Wendung und das Aufstehen in die Bewertung einbezogen.

1993 haben WELLS et al. statt dieser neunstufigen Bewertung ein fünfstufiges System beschrieben. Die Skala geht von 0 bis 4, wobei 0 für einen normalen Gang steht und 4 für Bewegungsunfähigkeit. Für die Bewertung wird die Symmetrie des Ganges begutachtet bzw. die grundsätzliche Möglichkeit aufzustehen.

SPRECHER et al. veröffentlichten 1997 ein Locomotion-Scoring-System, welches eine 5-stufige Bonitur beschreibt, bei der die Note 1 eine normale Bewegung darstellt und das Tier als gesund bewertet. Die Noten 2 – 5 spiegeln eine graduelle Verschlechterung der Körperhaltung wider. Die bei MANSON UND LEAVER (1988) beschriebenen Kriterien bezüglich der Schrittverkürzung und Gewichtsverteilung auf die Klauen wurden mit aufgenommen, jedoch wird das Hauptaugenmerk auf die Krümmung des Rückens gelegt. Diese wird sowohl im Stehen als auch beim Laufen bewertet (Tabelle 1).

Tabelle 1: Locomotion Scoring nach SPRECHER et al. (1997)

Lahmheits-Note	Klinische Beschreibung	Beurteilungskriterien
1	Normal	Die Kuh steht und geht mit einem geraden Rücken. Ihr Gang ist normal.
2	Leicht lahm	Die Kuh steht mit geradem Rücken, aber krümmt den Rücken beim Gehen. Ihr Gang bleibt normal.
3	Mittelmäßig lahm	Der Rücken ist sowohl im Stand als auch in der Bewegung gekrümmt. Der Gang ist verändert durch verkürzte Schritte mit einem oder mehreren Beinen.
4	Lahm	Der Rücken ist dauerhaft gekrümmt. Der Gang ist verzögert durch die Bevorzugung eines oder mehrerer Beine.
5	Schwer lahm	Die Kuh ist bewegungsunfähig oder kann ein oder mehrere Beine kaum belasten.

Diese Art des Vorgehens beim Locomotion Scoring wird laut einer Studie von SCHLAGETER-TELLO et al. (2014) am häufigsten in der Literatur beschrieben.

COOK (2003) hat die drei oben genannten Methoden für sich modifiziert. Es existieren nur vier Stufen, wobei 1 „nicht lahm“ ist, 2 „leicht lahm“, 3 „mittelmäßig lahm“ und 4 „schwer lahm“. Als klinisch lahm werden hier alle Kühe mit der Note 3 und 4 angesehen. Zusätzlich zur Rückenkrümmung, Schrittlänge und Gewichtsverteilung auf die Klauen wird bei dieser Bewertung auch die Laufgeschwindigkeit berücksichtigt.

GARBARINO et al. haben 2004 dagegen das Bewertungssystem nach SPRECHER et al. (1997) um eine Bewertungsnote erweitert. Die Skala geht von 0 bis 5. Die bei SPRECHER et al. (1997) als leicht lahm bezeichneten Kühe werden bei GARBARINO et al. (2004) als kaum lahm bezeichnet. Außerdem werden die leicht lahmen Kühe neu definiert. Diese haben zwar einen gekrümmten Rücken im Gehen und im Stehen, der Gang scheint aber weiterhin normal zu sein.

2006 haben AMORY et al. das System von SPRECHER et al. (1997) auf drei Bewertungsstufen reduziert. Es wird lediglich die Rückenkrümmung berücksichtigt und keine anderen Einflüsse wie die Schrittlänge oder Gewichtsverteilung auf die Klauen bewertet. Ein gerader Rücken im Gehen und im Stehen beschreibt ein gesundes Tier. Ist der Rücken im Stehen gerade, aber krümmt sich beim Laufen, erhält das Tier die Note 2. Ein im Gehen und im Stehen gekrümmter Rücken bekommt bei AMORY et al. (2006) die Note 3 und fasst die Noten 3, 4 und 5 von SPRECHER et al. (1997) zusammen.

FLOWER UND WEARY (2006) haben sich bei ihrer Lahmheitsbewertung an dem System von MANSON UND LEAVER (1988) orientiert. Es werden fünf Stufen beschrieben, wobei

die 1 einen flüssigen Bewegungsablauf darstellt und 5 eine stark eingeschränkte Bewegung. Die Ausgabe von halben Noten ist möglich, sofern ein Tier nicht alle Kriterien einer bestimmten Stufe erfüllt. Somit sind bei dieser Bonitur insgesamt 9 Stufen möglich. Im Gegensatz zu der Bewertung von MANSON UND LEAVER (1988) wird nicht die Ab- und Adduktion der Klauen betrachtet, sondern das Übertreten der Hinterklauen (tracking-up) und die Beweglichkeit der Gelenke. Außerdem wird zusätzlich zum gekrümmten Rücken nach SPRECHER et al. (1997) auch die Kopfbewegung beim Gehen beobachtet und bewertet.

BERRY et al. (2008) haben einen „Lameness:Speed Index“ entwickelt. Dieser greift für seine 5-stufige Lahmheitsbewertung auf die Beschreibungen von MANSON UND LEAVER (1988) und SPRECHER et al. (1997) zurück. Zusätzlich wird die Geschwindigkeit, mit der die Kuh den Weg entlangläuft, auf einer Skala von 1 bis 3 bewertet, wobei 3 für die langsamste Fortbewegung steht. Beide Bewertungen werden addiert und ergeben den Lameness:Speed Index. Dessen Notenskala reicht demnach von 2 bis 8.

THOMSEN et al. (2008) haben aus den bereits beschriebenen Beurteilungssystemen die klinischen Merkmale herausgesucht und diese optimiert in einer Tabelle zusammengefasst (Tabelle 2). Diese Methode soll eine einfachere Handhabung in der Praxis bieten.

Tabelle 2: Lahmheitsbeurteilungssystem nach THOMSEN et al. (2008)

Lahmheits-Kategorie	Unregelmäßiges Gangbild	Rücken im Gehen gekrümmt	Rücken im Stehen gekrümmt	Schrittlänge verkürzt	Kopfnicken	Betroffenes Bein erkennbar	Belastung des Beins wird vermieden
1 Normal	Nein	Nein	Nein	Nein	Nein	Nein	Nein
2 Unregelmäßiger Gang	Ja	Ja	Nein	(Ja) ¹	Nein	Nein	Nein
3 Leichte Lahmheit	Ja	Ja	Ja	Ja	Nein	Nein	Nein
4 Lahmheit	Ja	Ja	Ja	Ja	Ja	Ja	Nein
5 Schwere Lahmheit	Ja	Ja	Ja	Ja	Ja	Ja	Ja

¹Schrittverkürzung vielleicht erkennbar

2.3.3 Subjektivität der Beobachter

Das visuelle Locomotion Scoring kann auf zwei Arten durchgeführt werden. Zum einen live im Kuhstall, zum anderen anhand von Videoaufnahmen. Beide Vorgehensweisen haben ihre Vor- und Nachteile. Videoaufnahmen ermöglichen dem Beobachter einen Bewegungsablauf mehrfach anzusehen und komplexe oder sehr schnell ablaufende Situationen richtig zu bewerten. Andererseits zeigen Videoaufnahmen nur eine Momentaufnahme und es fehlt eventuell der Kontext, der bei der Live-Beurteilung einen Einfluss haben kann (BENCH et al. 1974). Bei visuellen Beurteilungen spielt es außerdem eine Rolle, ob ein Beobachter erfahren oder unerfahren ist. In Studien wird deshalb häufig die Interrater-Reabilität (Urteilerübereinstimmung) und die Intrarater-Reabilität (Wiederholbarkeit der Beurteilung eines Beobachters) der Beurteiler angegeben. Untersuchungen haben gezeigt, dass erfahrene Bewerber während des Live-Scorings eine geringere Inter- und Intrarater-Reabilität haben als bei Beurteilungen anhand von Videos. Unerfahrene Beobachter zeigen hingegen keine Unterschiede in der Beurteilung vom Video und live (SCHLAGETER-TELLO et al. 2015a). Außerdem wurden auch die verschiedenen Merkmale, die zur Beurteilung des visuellen Locomotion Scores beschrieben wurden, auf ihre Homogenität in der Beurteilung sowie in der Wiederholbarkeit untersucht. Dabei stellte sich heraus, dass die Merkmale Gewichtsverteilung auf die Beine, Asymmetrie des Ganges und Rückenkrümmung am besten mit dem Locomotion Score korrelieren (SCHLAGETER-TELLO et al. 2015b).

2.4 Verhaltens- und Leistungsänderungen durch Lahmheit

Die Aktivität von Kühen wird bereits in Form von Messungen der Stand- und Liegezeiten mithilfe von Pedometern in vielen Ställen erfasst. Diese können nicht nur der Brunsterkennung, sondern auch der Erkennung von Lahmheiten dienen. Lahme Kühe in Boxenlaufställen liegen signifikant länger pro Tag als gesunde Kühe (SINGH et al. 1993). Sie bewegen sich weniger und nehmen Haltungen ein, die ein Unwohlsein charakterisieren. Sowohl die Anzahl der Liegephasen pro Tag als auch die Länge der einzelnen Liegephasen erhöhen sich bei einer Lahmheit (ITO et al. 2010). Längere Liegezeiten gehen einher mit einer Verkürzung der Standzeiten. Außerdem reagieren lahme Kühe weniger auf den Anreiz frisch rangeschobenen Futters. Sie stehen vermehrt in den frühen Morgenstunden am Futtertisch; vermutlich, um Konflikte mit Artgenossen zu vermeiden (BLACKIE et al. 2011). Nach BACH et al. (2007) hingegen gehen Kühe mit

einem hohen Locomotion Score (LS) zur selben Zeit an den Futtertrog wie Kühe mit einem niedrigen LS. Jedoch bevorzugen lahme Kühe die Tröge, welche sich dicht am automatischen Melksystem befinden. Das gibt einen Hinweis auf einen verminderten Aktivitätsradius.

Bereits geringgradige Lahmheiten führen zu einer deutlichen Veränderung des Futteraufnahmeverhaltens bei Milchkühen. Kühe mit einem höheren LS verbringen weniger Zeit pro Tag mit der Futteraufnahme, fressen dabei aber deutlich schneller. Die kürzeste Zeit der Futteraufnahme zeigen Kühe in der ersten Laktation, die an einer Lahmheit leiden. Nichtsdestotrotz ist das Gewicht lahmer Kühe geringer als bei gesunden Kühen (NORRING et al. 2014).

Bei der Nutzung automatischer Melksysteme, welche selbständig und freiwillig von den Tieren aufzusuchen sind, zeigen lahme Kühe eine geringere Besuchsrate als gesunde Kühe (BORDERAS et al. 2008). Generell bewegen sich Kühe mit einem höheren LS langsamer als gesunde Kühe (CHAPINAL et al. 2010).

Grundsätzlich beginnen lahme Kühe zwar weniger aggressive Interaktionen, jedoch besteht kein Unterschied in der Anzahl aggressiver Gesten der Herdenmitglieder gegenüber lahmen Kühen. Ebenso konnte bei lahmen Kühen keine Veränderung in der Häufigkeit des Beleckens anderer Kühe festgestellt werden. Lahme Kühe selbst werden jedoch deutlich häufiger von anderen Kühen beleckt. Das Belecken anderer Kühe scheint demnach eine Geste zu sein, das Unbehagen anderer Herdenmitglieder, die Schmerzen erleiden, zu mildern (GALINDO UND BROOM 2002).

Lahmheiten wirken sich auch auf die Fruchtbarkeit von Milchkühen aus. In der Untersuchung von GARBARINO et al. (2004) zeigt sich, dass lahme Kühe eine 3,5-mal höhere Wahrscheinlichkeit haben, dass der Zyklus verspätet einsetzt.

Auch die Milchleistung leidet unter einem schlechten LS. Die Milchleistung pro Tag kann bis zu 1,5 kg verringert sein, wenn eine Kuh an einer Lahmheit leidet. Eine schwere Lahmheit im ersten Laktationsmonat kann die 305-Tage Leistung um bis zu 350 kg reduzieren (ARCHER et al. 2010).

2.5 Automatische Erfassung des Locomotion Scores

Es gibt drei Ansätze um einen Locomotion Score automatisch zu berechnen. Die *Kinetik* misst die wirkenden Kräfte im Raum unter Berücksichtigung der Masse. Die *Kinematik*

ist die Bewegungsgeometrie. Sie misst die Beschleunigung und den Weg der Gliedmaßen und berücksichtigt auch die Körperhaltung des Tieres. Die *indirekte Messung* wertet das Verhalten der Kühe und bestimmte Produktionsfaktoren aus. Im Folgenden werden ausgewählte Ansätze der einzelnen Methoden beschrieben.

2.5.1 Kinetische und kinematische Ansätze

Bereits Ende der 80er Jahre hat sich SCOTT (1989) mit Kühen und deren Gewichtsverteilung auf die einzelnen Gliedmaßen beschäftigt. Aus vorangegangenen Studien lagen bereits Kraft-Zeit-Kurven vor, welche von Kraftmessplatten gewonnen wurden. Diese zeigten, dass überwiegend Unterschiede zwischen den Vorder- und Hinterbeinen bestehen, nicht aber bei den kontralateralen Gliedmaßen. Aufgrund der Tatsache, dass lahme Kühe die schmerzende Gliedmaße weniger belasten, wurde überprüft, ob dies auf der Kraftmessplatte sichtbar ist. In den Messkurven sind die vertikalen und horizontalen Kräfte, die beim Auftreten entstehen, sichtbar. Die vertikalen Kräfte lassen sich unterteilen in eine Abbrems- und Beschleunigungsphase. Diese Kraft-Zeit-Kurven zeigten bei gesunden Kühen gleichgroße Kräfte beim Abbremsen und beim Beschleunigen. Bei einigen lahmen Kühen war die vertikale Kraft, die vom lahmen Bein ausgeübt wird, vermindert und die horizontalen Kräfte nicht mehr ausgeglichen. Es wurden auch signifikante Unterschiede zwischen dem lahmen und dem kontralateralen Bein festgestellt. SCOTT (1989) konnte anhand seiner vorliegenden Daten allerdings keine Verallgemeinerungen für alle lahmen Tiere machen. Der Gang der Kühe variiert und Veränderungen in der Bewegung, verursacht durch Lahmheit, sind sehr unterschiedlich und könnten abhängig vom jeweils erfahrenen Schmerz sein.

FLOWER et al. (2005) waren die ersten, die sich mit der kinematischen Technik zur Lahmheitsbeurteilung beschäftigt haben. Diese verspricht eine geringe Beanspruchung humaner Ressourcen. Ihre Hypothese war, dass Klauenerkrankungen (insbesondere Klauensohlengeschwüre) den Gang der Kuh beeinflussen, was kinematisch messbar sein soll. Es zeigte sich, dass Kühe mit einem Klauensohlengeschwür gegenüber gesunden Kühen kürzere und flachere Schritte machen und eine längere Schwungphase haben. Die gesamte Zeit für einen Schritt ist erhöht.

RAJKONDAWAR et al. (2006) haben die Idee von SCOTT (1989) aufgegriffen und weiterentwickelt. Auch sie haben die vertikalen und horizontalen Kräfte einer Kuh beim

Auftreten gemessen. Die Bodenplatte bestand aber aus zwei nebeneinanderliegenden Platten. Durch eine Einfassung der Platten in einen Rahmen wurden die rechten Gliedmaßen der Kuh auf die rechte Platte und die linken auf die linke Platte geführt. Anhand der vertikalen Kraftmessung wurde eine Beinbewegungsvariable (limb movement variable = LMV) ermittelt. Die Hypothese war, dass anhand von fünf Beinbewegungsvariablen eine Lahmheit erkannt werden kann. Damit der Kuhverkehr nicht durch eine Tiervereinzelung aufgehalten wird, haben TASCH UND RAJKONDAWAR (2004) die Software SoftSeperatorTM entwickelt. Diese zeichnet die Datensätze von mehreren Kühen, die in einer Gruppe hintereinander über die Messplatte laufen, auf und legt sie mit tierindividuellen Nummern ab. In den Auswertungen zeigte der Schweregrad der Klauenläsion die engste Korrelation mit der Beinbewegungsvariable (RAJKONDAWAR et al. 2006).

Diese von RAJKONDAWAR et al. (2006) beschriebene automatische Lahmheitserkennung wird unter dem Namen StepMetrixTM von BouMatic kommerziell vertrieben. 2007 haben BICALHO et al. die Genauigkeit dieses automatischen Systems mit der des visuellen Locomotion Scorings verglichen. In einer ersten Studie wurde untersucht, ob ein Locomotion Score mit einer Skala von 1 bis 5 ab der Note 3 oder ab der Note 4 als „lahm“ definiert werden sollte. Dafür wurden die Ergebnisse des Klauenschnitts und des visuellen Scorings gegenübergestellt. Der Korrelationskoeffizient nach Spearman lag bei einer Einstufung als lahm bei Kühen mit einem visuellen Locomotion Score (vLS) von ≥ 3 bei $0,48 \pm 0,04$ und bei Kühen mit einem vLS ≥ 4 bei $0,41 \pm 0,04$. In der zweiten Studie lag die Sensitivität der „lahm = ≥ 3 -Kühe“ bei 67,5 % und die Spezifität bei 84,6 %. Für das automatische System wurden Kühe ab dem Wert 39 als lahm bewertet. Anhand dieser Einstufung ergaben sich eine Sensitivität von 22,2 % und eine Spezifität von 93,8 %, verglichen mit den Ergebnissen des Klauenschnitts.

PASTELL UND KUJALA (2007) haben Untersuchungen zur Gewichtsverteilung von Milchkühen während des Melkens im Melkroboter durchgeführt. Dazu wurden im Melkroboter vier Messplatten installiert, welche das Gewicht jedes Beins dynamisch messen können. Datensätze von Tieren, die nicht mit allen vier Beinen richtig auf den Platten standen, wurden durch einen bestimmten Algorithmus entfernt. Das System war in der Lage, die durchschnittliche Gewichtsbelastung der einzelnen Beine zu bestimmen und die Tritte während des Melkens zu zählen. Mit Hilfe eines neuronalen Wahrscheinlichkeitsnetzwerks (PNN = probabilistic neural network) gelang es, 96,2 %

der Messungen korrekt der Kategorie lahm oder gesund zuzuordnen und 100 % der lahmen Kühe aus der visuellen Kontrolle wurden als solche vom System identifiziert. Der Anteil der Messungen, die Fehlalarme waren, beschränkte sich auf 1,1 %. In dem Artikel von 2008 haben PASTELL et al. ihre weiterführenden Versuche erläutert. Dabei haben sie eine Matte mit einem integrierten elektromechanischen Film der Firma Emfit getestet. Der Vorteil dieser Matte ist, dass sie nicht auf die Nutzung im Melkroboter beschränkt ist, sondern in jedem Laufgang genutzt werden kann, denn sie misst ausschließlich dynamische Kräfte. Es wurden zwei Sensorstreifen mit den Maßen 150 x 58 cm parallel auf den Boden und zum Schutz eine 15 mm dicke Gummimatte darübergelegt. Auch auf dieser Matte waren Lahmheiten, durch die geringere Kraft, die auf das lahme Bein übertragen wird und die verlängerte Zeit, die es den Boden berührt, erkennbar. Besonders in Kombination mit den zuvor entwickelten Bodenplatten im Melkroboter schien dieses Studienergebnis vielversprechend.

MAERTENS et al. (2011) haben die Entwicklung von TASCH UND RAJKONDAWAR (2004) aufgegriffen und weiterentwickelt. Sie haben eine druckempfindliche Matte entwickelt, die insgesamt 20 verschiedene kinematische Variablen misst. Dieses System erreichte eine Sensitivität von 76 % – 90 % und eine Spezifität von 86 % – 100 %. Bei den Versuchen zeigten sich die Variablen für Asymmetrie und Laufgeschwindigkeit als sehr vielversprechend.

Da sich die Druckmessplatten aufgrund der hohen Anschaffungskosten und der schweren Integration in die praktischen Betriebsabläufe nicht durchsetzen konnte, untersuchten VAN DE GUCHT et al. (2017) mögliche Einsparungen innerhalb des Systems, um dieses attraktiver für die Praxis zu machen. Dazu wurde die Länge der Matte gekürzt und die Auflösung der Sensoren minimiert. Es hat sich gezeigt, dass statt der 4,88 m auch eine Länge von 3,28 m ausreichend ist, um eine Schrittfolge aufzunehmen. Die Größe eines Sensorelements in der Matte konnte um das Vierfache vergrößert werden. Trotz dieser Einsparungen blieb die Gesamtgenauigkeit bei 58 %.

2.5.2 Lahmheitsbeurteilung anhand von Videoanalysen

BAHR et al. (2008) haben anhand von Videoaufnahmen Parameter aus den Bildern analysiert, welche mit den visuellen Scores von Beobachtern korrelieren. Dabei wurden das Übertreten der hinteren Gliedmaßen und die Krümmung des Rückens untersucht. Es konnte eine positive Beziehung zwischen dem Übertreten der Hinterbeine und dem

visuellen Locomotion Score gezeigt werden. Auch die stärkere Wölbung des Rückens mit ansteigendem visuellen Score und damit verbunden der kleinere Radius konnten nachgewiesen werden. SONG et al. (2008) hatten den Anspruch, die von BAHR et al. (2008) beschriebenen Messungen zum Übertreten zu automatisieren. Die Videos wurden als Sequenzen von Bitmaps wiedergegeben. Wie bereits von anderen Autoren beschrieben (TELEZHENKO et al. 2005, ESPEJO et al. 2006, FLOWER UND WEARY 2006), korreliert ein positiver Übertritt-Wert stark mit einem hohen Locomotion Score. Auch SONG et al. (2008) konnten diese Korrelation nachweisen und den Übertritt-Wert automatisch aus den Videos berechnen. Verglichen mit der manuellen Berechnung der Übertritt-Werte zeigte sich eine Korrelation von 94,83 % und ebenso eine positive, lineare Beziehung zu den visuellen Locomotion Noten.

POURSABERI et al. (2010) haben versucht mit Hilfe von Bildverarbeitungstechnik einen Algorithmus zu entwickeln, der Lahmheiten anhand der Rückenkrümmung erkennt. Dabei wurden Videos von der Seite aufgezeichnet und an die Rückenlinie der Kuh ein Kreis approximiert. Der Radius des Kreises war der Messwert, der die Rückenkrümmung widerspiegelt. 184 Videos wurden ausgewertet und nur sechs Fehlklassifizierungen entstanden. Der Abstand der Fehlklassifizierungen betrug maximal einen Score. Es wurden nur gesunde Kühe fälschlicher Weise als lahm ausgegeben, nicht aber lahme Kühe als gesund angezeigt. Somit lag die Sensitivität bei 100 % und die Spezifität bei 97,6 %.

2013 haben VIAZZI et al. die Methode von POURSABERI et al. (2010) aufgegriffen und weiter untersucht. In den Versuchen hatten sie acht Kühe, die in der gesamten Versuchszeit alle Lahmheitsausprägungen zeigten (Score 1-3). Bei den Auswertungen wurde deutlich, dass die Messwerte der Kühe sich in den unterschiedlichen Scores unterscheiden. Das bedeutet, dass feste Grenzwerte für alle Tiere nicht zielführend sind. Durch tierindividuelle Grenzwerte konnte die Genauigkeit um 10 % erhöht werden.

2014 haben VIAZZI et al. das erste Mal eine 3D Kamera zur Lahmheitsbeurteilung genutzt. Dabei wurde die Rückenlinie nachgebildet, in zwei Teile ausgehend vom höchsten Punkt des Rückens geteilt und eine bestimmte Höhe L_1 (siehe Abbildung 3) berechnet. Es gelang eine Spezifität von 91 % und eine Sensitivität von 82 % zu erreichen.

VAN HERTEM et al. (2014) haben mit derselben Technik und im selben Versuchsbetrieb wie VIAZZI et al. (2014) weitere Versuche mit der 3D Kamera durchgeführt. An vier

aufeinander folgenden Tagen wurden sowohl automatische Scores berechnet als auch live visuelle Locomotion Scores vergeben. Anhand dieser Daten wurden drei verschiedene Klassifizierungsmethoden getestet. Wurden die Klassengrenzen so gelegt, dass möglichst viele Kühe richtig klassifiziert wurden, lag die Übereinstimmung bei 53 % für alle 744 Messungen. Die Nutzung von kuhindividuellen Daten über vier Messungen führte zu einer höheren Übereinstimmung. Die höchste Übereinstimmung wurde erreicht, wenn statt der 5-stufigen Skala eine zweistufige Klassifikation in lahm und gesund stattfand. In diesem Fall wurde eine Genauigkeit von 81 % erreicht.

JABBAR et al. (2017) haben die Nutzung von 3D Daten weiterentwickelt. Sie haben sowohl die Wirbelsäule als auch die Hüftknochen in den 3D Bildern lokalisiert und Veränderungen während des Ganges gemessen. Dadurch konnte die Gleichmäßigkeit des Ganges ermittelt werden und eine besonders frühe Erkennung einer Lahmheit ist möglich. Die Einstufung erfolgte in Scores von 1 – 5, wobei besonders die Grenze zwischen dem ersten und zweiten Score wichtig war, da an dieser Stelle eine beginnende Lahmheit sichtbar wird. Mithilfe dieses automatischen Systems sollte es dem Landwirt möglich sein, eine Kuh im frühestmöglichen Stadium zu untersuchen und mögliche Folgen aufgrund einer Lahmheit zu verhindern. Durch die Nutzung einer linearen Support-Vektor-Maschine konnte bei einer zweistufigen Klassifizierung eine Genauigkeit von 95,7 %, eine Sensitivität von 100 % und eine Spezifität von 75 % erreicht werden.

PLUK et al. (2012) haben die Vorteile einer Druckmessmatte und von Videoaufnahmen vereint. Anhand der Druckmessmatte konnte der Moment des Auftretens und des Abfußens ermittelt werden und in den Bildaufnahmen in diesen Momenten der Winkel der jeweiligen Beine gemessen werden.

BLACKIE et al. (2013) haben mit Videoaufnahmen von der Seite versucht, anhand der Schrittlänge und anderer Merkmale für die Lahmheitsbeurteilung einen Score zu entwickeln und automatische Messungen durchzuführen. Für die Auswertung der Videoaufnahmen waren Markierungen an den Kuhbeinen notwendig, welche das System weniger praxisrelevant scheinen lassen.

2.5.3 Indirekte Messungen

Indirekte Methoden zur Messung von Lahmheiten sind beispielsweise die Messung des Liegeverhaltens anhand von elektronischen Datenloggern. ITO et al. (2010) kamen

jedoch zu dem Schluss, dass das Liegeverhalten allein nicht für das Lahmheitsmonitoring nutzbar ist. Es könne lediglich als ein nützliches, unterstützendes Instrument eingesetzt werden.

Auch BLACKIE et al. (2011) haben Aktivitätsmessungen vorgenommen. Sie konnten eine signifikant längere Liegezeit bei Kühen mit einem Locomotion Score von 3 nachweisen als bei Kühen mit einem Score von 1 oder 2. Darüber hinaus haben die Kühe mit einem Locomotion Score von 3 acht Liter weniger Milch produziert als die Kühe mit einem Score von 1.

ALSAAOD et al. (2012) haben Pedometer ausgewertet, um Lahmheiten aufzuspüren. Dabei konnten große Unterschiede im Verhalten zwischen den Kühen festgestellt werden. Diese Unterschiede waren deutlich größer als die des Verhaltens einer bestimmten Kuh im lahmen und im gesunden Zustand. Deshalb wurde anstelle von absoluten Grenzwerten die Abweichung vom normalen Verhalten als Merkmal untersucht. Dieses erreichte eine Genauigkeit von 76 %.

2013 haben DE MOL et al. aufgenommene Bewegungsdaten mit Produktionsdaten wie der Milchmenge und Kraftfutteraufnahme pro Tag zusammen ausgewertet und konnten für die Lahmheitsbeurteilung eine Sensitivität von 85,5 % und eine Spezifität von 88,8 % erreichen.

2.5.4 Versuchsaufbau und Messtechnik zur automatischen Erfassung des Locomotion Scores

In diesem Abschnitt wird der Versuchsaufbau von ausgewählten, veröffentlichten Untersuchungen beschrieben und auf die verwendete Messtechnik eingegangen, um diese mit den in der vorliegenden Arbeit beschriebenen Prozessentwicklungen vergleichen zu können.

Bei SCOTT (1989) wurden vier lahme Kühe als Datengrundlage gewählt, um die Funktion der Druckmessplatten zu überprüfen. Dieser Stichprobe wurden acht gesunde Kühe gegenübergestellt, welche als Kontrollgruppe dienten. Diese Kühe gingen selbstständig über die Messplatte und wurden nicht geführt, wie es in vorherigen Untersuchungen beschrieben wurde. Es wurde bereits vermutet, dass das Führen der Tiere zu einer Beeinflussung ihres Ganges führt.

Für die Studie von RAJKONDAWAR et al. (2006) wurde allen Kühen vor Versuchsbeginn von einem Veterinär die Klauen geschnitten und alle Läsionen dokumentiert und benotet. Auch der Gang wurde beurteilt und jedem Bein eine eigene Note gegeben. Nur wenn alle vier Beine die Note 1 erhielten, wurde die Kuh als gesund eingestuft. Das automatische System wurde im Laufgang auf dem Rückweg vom Melkstand zum Stall angebracht und konnte von den Kühen zwei bis drei Stunden vor dem Melken aufgesucht werden. Die Daten wurden innerhalb von sechs Monaten von 21 Kühen mit unterschiedlichem Alter aufgenommen. Jeden Tag wurden für jedes Bein fünf Beinbewegungsvariablen berechnet. Anhand von 700 Auswertungen von Hinterbeinmessungen wurde ein gangbasiertes, ein läsionsbasiertes und ein kombiniertes Modell aus Gang und Läsion entwickelt und in einer logistischen Regression getestet.

BICALHO et al. (2007) haben zwei Studien durchgeführt. In der ersten Studie wurde das visuelle Locomotion Scoring einmal pro Woche durch einen Veterinär durchgeführt, wenn die Kühe den Melkstand verließen. Parallel wurde eine klinische Studie durchgeführt, die den Effekt einer Behandlung einer lahmen Kuh analysiert. Dafür wurde zu Beginn der Versuche allen Kühen zufällig die Nummer 1 oder 2 gegeben. Alle Kühe, welche die 1 bekommen haben, wurden bei einem visuellen Score von 4 oder 5 als lahm eingestuft und behandelt, alle Kühe mit einer 2 dagegen bei einem visuellen Score von 3, 4 oder 5. So wurde nur ein gewisser Anteil der Kühe mit einem Score von 3 auch behandelt. Insgesamt wurden 610 Kühe in den Versuch aufgenommen. Davon wurden 246 aufgrund der visuellen Einstufung ($LS \geq 3$) vom Klauenschneider behandelt und 213, weil der routinemäßige Klauenschnitt an der Reihe war. Der Klauenschnitt wurde immer vom selben Veterinär durchgeführt. Nur die erste Behandlung jeder Kuh wurde für die Auswertung herangezogen. Der Aufbau der zweiten Studie war ähnlich. Hier wurde das visuelle Scoring von drei Veterinären durchgeführt, die zuvor gemeinsam für die Durchführung geschult wurden. Außerdem war das StepMetrixTM-System im Rücktreibgang zum Stall aufgestellt. Dieses wurde von den Kühen dreimal pro Tag nach dem Melken durchlaufen. Das System hat die Mittelwerte der Tagesdaten des StepMetrixTM-Systems für die letzten sieben Tage gemittelt und den automatischen Score für denselben Tag, an dem das visuelle Locomotion Scoring durchgeführt wurde, ausgegeben. Der automatische Score hat eine Skala von 1 bis 100, wobei eine Kuh ab dem Wert 39 als lahm ausgegeben wurde. In Studie 2 wurden insgesamt 518 Kühe ausgewertet.

Bei PASTELL et al. (2008) wurden alle Messungen von einer digitalen Videokamera aufgenommen, welche nicht genauer spezifiziert wurde.

FLOWER et al. (2005) haben eine Videokamera eingesetzt, welche 60 Bilder pro Sekunde aufnimmt. Die Kamera wurde in 6,75 m Entfernung aufgestellt und hat die Kühe über eine Strecke von 7,05 m von der linken Seite aufgenommen. Über der Kamera wurde eine 100 Watt Lampe montiert, welche den Bildausschnitt beleuchtet. Die Kühe wurden mit reflektierenden Bändern an jedem Bein markiert. Insgesamt bestand die Testherde aus 48 Kühen. Diese wurden in zwei Gruppen á 24 Kühe aufgeteilt, die jeweils an sieben aufeinanderfolgenden Tagen morgens nach dem Melken aufgenommen wurden. Die Videoaufnahmen zeigen mindestens zwei aufeinanderfolgende Schritte jeder Kuh. Aufnahmen, auf denen die Kühe angehalten haben, gestolpert oder gerutscht sind, gekotet oder uriniert haben oder andere unnormale Bewegungen gezeigt haben, wurden gelöscht. Auch Aufnahmen, bei denen die Kühe zu dicht hintereinander laufen oder wenn ein Reflektorband fehlt, wurden nicht berücksichtigt. Acht bis neun Wochen nach den Videoaufzeichnungen wurden die Testtiere von einem professionellen Klauenschneider untersucht und alle Läsionen dokumentiert. Auf diese Weise sollte herausgefunden werden, ob zur Zeit der Videoaufnahmen bereits eine Klauenerkrankung vorlag, welche erst acht Wochen später beim Klauenschneider sichtbar wird. In die endgültige Auswertung gingen 38 Kühe ein.

Neun Monate nach den ersten Videoaufnahmen haben FLOWER et al. (2005) an 26 Kühen drei zusätzliche morphometrische Messungen durchgeführt. Gemessen wurden die Höhe der Schulter (T3), die Höhe des Steißbeins (Cc1) und die Länge des Rückens (Cc1-T3). Alle Messungen wurden von jeder Kuh dreimal aufgenommen, es konnten aber keine Unterschiede der Messergebnisse zwischen gesunden Kühen und solchen mit Klauensohlengeschwüren beobachtet werden.

Auch BAHR et al. (2008) haben Videos von Kühen beim Gehen aufgenommen. Die Kamera wurde in 8 m Entfernung und 1,5 m Höhe aufgestellt und hat Videos mit 680 x 480 Pixeln und 30 Bildern pro Sekunde produziert. Insgesamt wurden Daten von zehn Kühen aufgenommen. Für die Berechnung des Übertretens der Hinterbeine mussten die Koordinaten, an denen die Klauen den Boden berühren, manuell auf jedem Bild markiert werden. Daraus konnte dann automatisch der Abstand der Position des Hinterfußes zur Stelle, an welche der Vorderfuß stand, berechnet werden.

Auch für die Berechnung der Rückenkrümmung mussten zunächst manuell die Hüfte, die Schulter und der Mittelpunkt auf dem Rücken zwischen diesen beiden Punkten markiert werden. Anschließend wurde der Radius eines Kreises, der alle drei Punkte berührt, berechnet. Um die Änderung während des Gehens zu berücksichtigen, wurde die Markierung der Punkte und Berechnung der Krümmung an vier Bildern pro Kuh durchgeführt, jeweils während der Belastung eines Beines. Als Endwert wurde der Mittelwert der Radien ausgegeben.

SONG et al. (2008) haben versucht, die Berechnung des Übertritts zu automatisieren. Auch hier wurden die Videos von der Seite aus 6 m Entfernung aufgenommen. Die Kühe haben einen 6 m langen und 0,9 m breiten Weg, begrenzt durch ein Metallgitter, außerhalb des Kuhstalls einzeln und ohne Zwang passiert. Die entstandenen Videos hatten eine Auflösung von 1024 x 768 Pixel mit 15 Bildern pro Sekunde. Um die Kuhkörper aus den Bildern zu erkennen, wurde das leere Hintergrundbild von dem Bild mit der Kuh subtrahiert. So konnte die Kuh erkannt werden. Anschließend wurden die Klauen in den Bildern ermittelt und der Übertritt berechnet.

POURSABERI et al. (2010) hatten zwei Testherden. Die erste Stichprobe umfasste die Videos von 28 Kühen, die aufgezeichnet wurden, während sie von verschiedenen Beobachtern visuell eingestuft wurden. Dabei gingen die Kühe über einen 1,2 m breiten und 6 m langen Betongang und die Kamera war seitlich in 8 m Entfernung und einer Höhe von 1,5 m aufgestellt. Die Auflösung der Bilder war 640 x 480 Pixel und es wurden 30 Bilder pro Sekunde aufgezeichnet. In der zweiten Stichprobe wurden 66 Kühe aufgenommen. Die Aufnahmen fanden einmal pro Woche an drei aufeinanderfolgenden Wochen statt. Dazu liefen die Kühe einen 1 m breiten und 6 m langen Gang entlang, der mit einer modifizierten Druckmessmatte ausgelegt war. Eine Kamera wurde in 2,5 m Höhe und 3,5 m Entfernung angebracht, die 30 Bilder pro Sekunde mit einer Auflösung von 1024 x 768 Pixel aufnahm. Insgesamt standen von der zweiten Testherde 156 Videos zur Verfügung. Die Videos wurden in JPEG Bilder umgewandelt und für die Auswertung wurde jedes fünfte Bild genutzt (sechs Bilder pro Sekunde statt 30). Das leere Hintergrundbild wurde mit den Kuhaufnahmen verglichen und die Position der Kuh auf dem Bild automatisch mit einem Algorithmus ermittelt (Abbildung 2). Anschließend wurde ein weiterer Algorithmus angewandt, der den Bildausschnitt mit dem Kuhkörper aus dem Bild extrahiert. Da sich teilweise die Umrisse der Kuh nur schwer vom Hintergrund abgehoben haben, zeigte das

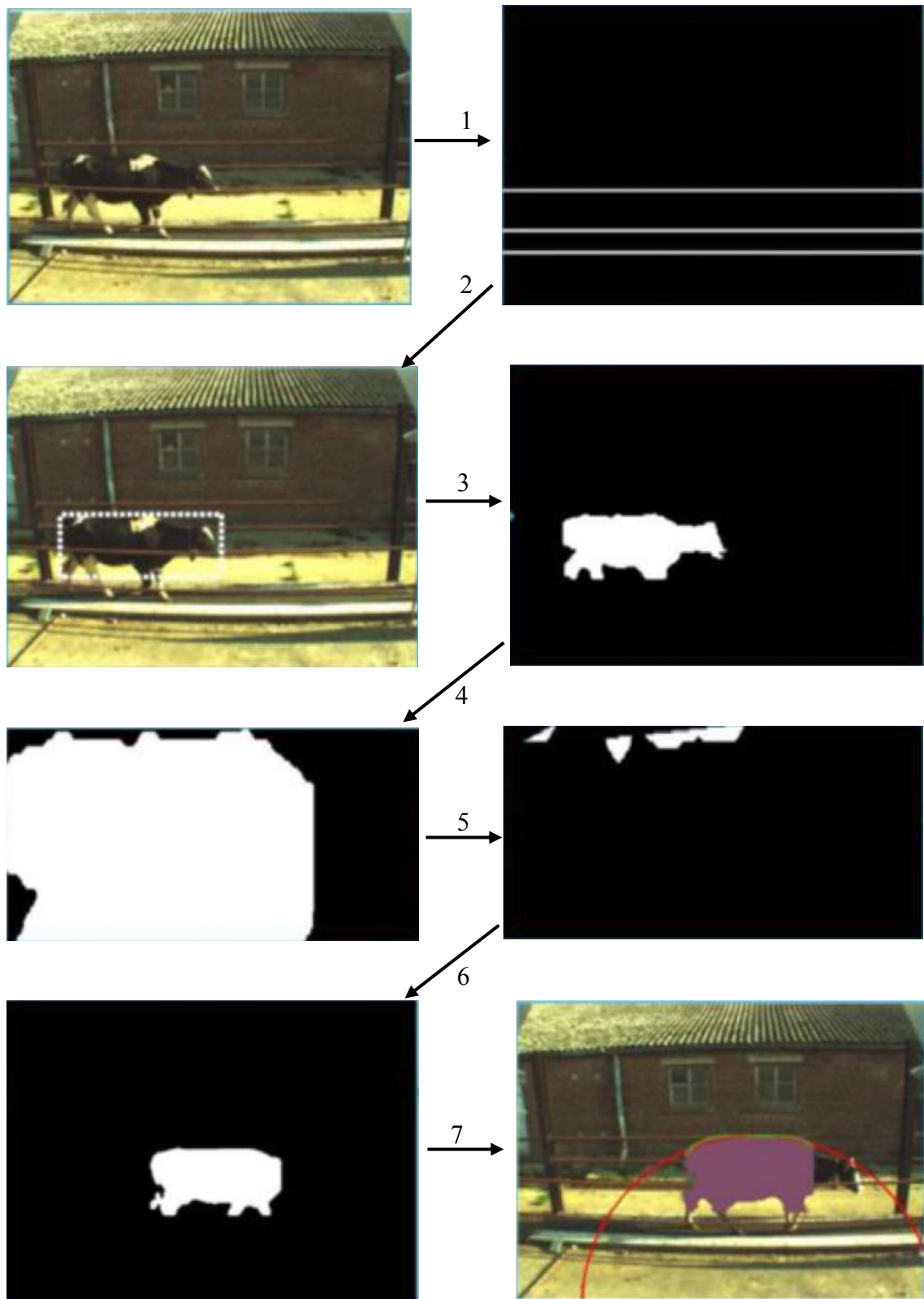


Abbildung 2: Flussdiagramm der Methode von POURSAHERI et al. (2010). Input des Systems sind Videoaufzeichnungen und es wird ein numerischer Wert ausgegeben, der im Verhältnis zur Rückenkrümmung steht. Dafür wird die Kuh vom Hintergrund separiert und anhand des Binärbildes die Rückenkrümmung berechnet.

Differenzbild nie den genauen Verlauf der Rückenlinie. Erschwerend kamen Veränderungen im Hintergrund hinzu, die schon durch veränderte Lichtverhältnisse entstanden. Das Bild wurde in ein Binärbild überführt, welches nur aus schwarzen und weißen Pixeln besteht. Die weißen Pixel stellten die Kuh dar. Fehlstellen innerhalb des Kuhumrisses wurden automatisch aufgefüllt und Artefakte extrahiert, da das größte zusammenhängende weiße Feld als Kuh definiert wurde. Im nächsten Schritt wurden der Hals und der Kopf der Kuh ausgeblendet, indem die Summe der weißen Pixel auf der X-Achse des Bildes mit einem Grenzwert versehen wurde, denn der Hals und der Kopf enthalten weniger Pixel als der Kuhkörper. Abschließend wurden Lücken in der Rückenlinie aufgefüllt. Für die Berechnung der Rückenkrümmung wurden drei Punkte auf der Rückenkontur markiert. Die gesamte Länge der Oberlinie wurde berechnet. 5 % vom ersten und vom letzten Punkt entfernt wurden zwei Punkte gesetzt. Der Dritte Punkt markierte die Mitte zwischen diesen Punkten. Anschließend wurde ein Kreis, welcher alle drei Punkte berührt, erstellt. Der Radius (r) dieses Kreises wurde berechnet und die Rückenkrümmung ausgegeben als $1/r$.

Da Kühe das Gewicht vom lahmen Bein auf das gesunde verlagern und Lahmheiten überwiegend in den Hinterbeinen auftreten, wurde sich auf die Bilder konzentriert, in welchen ein Hinterbein den Boden berührte. Vier Bilder jedes Videos, auf denen ein Bein den Boden berührte, wurden automatisch durch einen Algorithmus ausgewählt. Auf zweien berührte das linke Hinterbein den Boden und auf zwei anderen das rechte Hinterbein. Der Durchschnitt der berechneten Radien der vier Bilder wurde der jeweiligen Kuh zugeordnet. Die Messwerte wurden den visuellen Scores der Beobachter gegenübergestellt und so Grenzwerte für die automatisch ausgegebenen Scores ermittelt.

PLUK et al. (2012) haben eine Kamera in 2,5 m Höhe und 3,5 m Entfernung vom Laufgang installiert. Diese hat 20 Bilder pro Sekunde mit 1024 x 768 Pixeln aufgenommen. Die Kamera war ausgerichtet auf einen 1 m breiten und 6 m langen Gang außerhalb des Stallgebäudes, dessen Boden mit einer Druckmessmatte ausgelegt war und als Hintergrund ein grüner Vorhang aufgehängt wurde, um einen besseren Kontrast zwischen Vorder- und Hintergrund zu erreichen. Damit die Tiere einzeln den Versuchsaufbau durchlaufen, wurde davor ein Selektionstor angebracht, welches automatisch die Transpondernummer der Kuh ausliest und automatisch für 15 Sekunden die Aufnahme der Kamera und der Druckmatte startet. Durch denselben Zeitstempel waren die Daten einander zuzuordnen. Die automatische Datenaufnahme erfolgte über

fünf Wochen. Von 504 Datensätzen konnten 104 nicht ausgewertet werden, weil die Synchronisation zwischen Matte und Kamera nicht gegeben war, die Lichtverhältnisse ungünstig waren oder die Kuh nicht gleichmäßig gelaufen ist.

VIAZZI et al. (2014) haben die 3D Kinect Kamera genutzt. Diese zeichnet 30 Bilder pro Sekunde auf und erzeugt Tiefenbilder mit 640 x 480 Pixel mit einer Auflösung von 1 cm aus 2 m Entfernung. Da der Sensor sehr empfindlich gegenüber Sonnenlicht ist, wurden die Versuche während der Nacht durchgeführt. Außerdem wurde eine 2D Kamera installiert, welche die Kühe von der Seite aufnimmt. Der 3D Sensor wurde 3,15 m über dem Boden befestigt und die 2D Kamera in 6 m Entfernung vom Korridor. Zur Lahmheitsbeurteilung wurden zwei Algorithmen verwendet, einer für die 2D Kamera und einer für den 3D Sensor. Bei der 2D Kamera wurde manuell das Bild ausgesucht, bei dem gerade der Hinterfuß den Boden berührt. Anschließend wurde manuell die Rückenlinie nachgezeichnet und ausgehend vom höchsten Punkt in zwei Hälften geteilt. An beide Rückenhälften wurden Ellipsen angepasst. Für die 3D Aufnahmen musste zunächst ein Algorithmus entwickelt werden, der die einzelnen Kühe aus den 3D Aufnahmen erkennt, da keine automatische Erkennung bzw. kein automatisches Start-Stopp-Signal gegeben werden konnte. Außerdem wurden nur Bilder zur Auswertung genutzt, auf denen der gesamte Kuhkörper sichtbar ist. Dafür wurden alle Bilder entfernt, die weniger als 8000 Pixel (85 % der durchschnittlichen Kuhgröße in Pixel) aufwiesen. Im nächsten Schritt wurde die Kontur der Kuh berechnet und das Binärbild genutzt, um den Kopf der Kuh aus dem Bild auszublenden. Der höchste Punkt in der Nähe der Orientierungsachse (10 % der Kuhbreite) zeigte die Wirbelsäule. Mit einer Polyfit Funktion (Abbildung 3) wurde die Rückenlinie nachgezeichnet. Dabei wurden vom höchsten Punkt der Rückenlinie (R) zwei Ellipsen gezeichnet und deren Orientierung φ_1 und φ_2 berechnet. Der Schnittpunkt der zwei kleineren Achsen der Ellipsen wurde bestimmt und der resultierende Winkel φ_3 berechnet. L_1 gab den vertikalen Abstand zwischen diesem Schnittpunkt und dem höchsten Punkt (R) wider. Der Durchschnitt jedes Parameters, der in den Einzelbildern (\varnothing 2,5 Bilder/Kuh) berechnet wurde, wurde für die Bestimmung des Locomotion Scores herangezogen. Zur Klassifizierung wurde ein Entscheidungsbaum genutzt, der alle vier Parameter einbezieht.

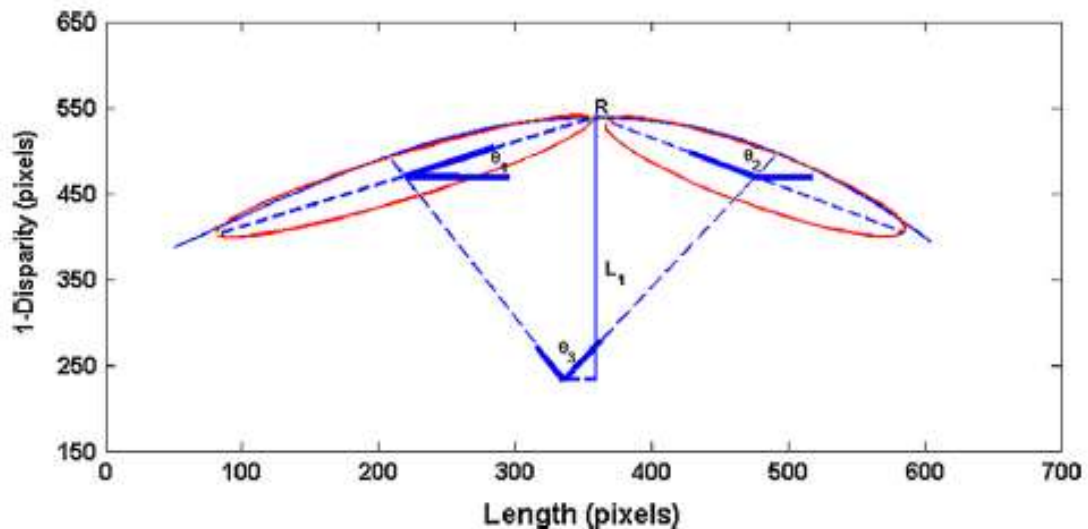


Abbildung 3: Parameter θ_1 , θ_2 , θ_3 und L_1 einer nachgebildeten Rückenlinie einer Kuh mittels 3D Sensor (VIAZZI et al. 2014)

VAN HERTEM et al. (2014) haben mit derselben Technik und im selben Versuchsbetrieb wie VIAZZI et al. (2014) weitere Versuche mit dem 3D Sensor durchgeführt. Dabei wurde der Sensor allerdings 5 cm höher montiert und die 2D Kameraaufnahmen nur zur Tiererkennung genutzt. Außerdem wurde 0,5 m vor dem Videosichtfeld eine Lichtschranke eingesetzt, um die Videoaufnahmen zu triggern. Eine Aufnahme dauerte 4 Sekunden. In dieser Versuchsreihe wurde zur Berechnung der Rückenkrümmung der Median aller Bilder, auf denen die Kuh vollständig abgebildet ist ($\approx 4,9$ Bilder pro Kuh), verwendet.

JABBAR et al. (2017) haben die ASUS xTion Pro Live für die 3D Aufnahmen genutzt. Die Tiererkennung erfolgte anhand der manuell dokumentierten, visuell erkannten Brandnummern der Tiere. Diese manuelle Dokumentation wurde anschließend mit den Zeitstempeln einer radio-frequency identification (RFID) zur Tiererkennung verglichen, um die 3D Bilder den Kühen zuzuordnen. Der 3D Sensor wurde in einer Höhe von 3,69 m installiert, sodass der Laufgang über 6 m aufgezeichnet wurde. Dadurch war es möglich, zwei vollständige Schrittfolgen zu analysieren. Durchschnittlich standen 70 Bilder zur Auswertung zur Verfügung. Von jeder Kuh in der Studie standen mindestens drei Scores aus einem Zeitraum von drei Wochen zur Verfügung. Hier wurden nur Kühe mit einem $vLS \leq 3$ ausgewertet. Es wurde ein Tiefenbild erstellt und Höhenmessungen durchgeführt. Der genutzte Algorithmus war in der Lage, stark abgerundete Körperpartien, wie die Hüfthöcker und die Wirbelsäule, zu segmentieren. Abbildung 4 zeigt den Prozess der Erkennung vom rohen Tiefenbild über die

Subtraktion der Umgebung zur Darstellung der höchsten Körperpartien. Dieser Prozess wurde für jedes Bild eines Datensatzes wiederholt.

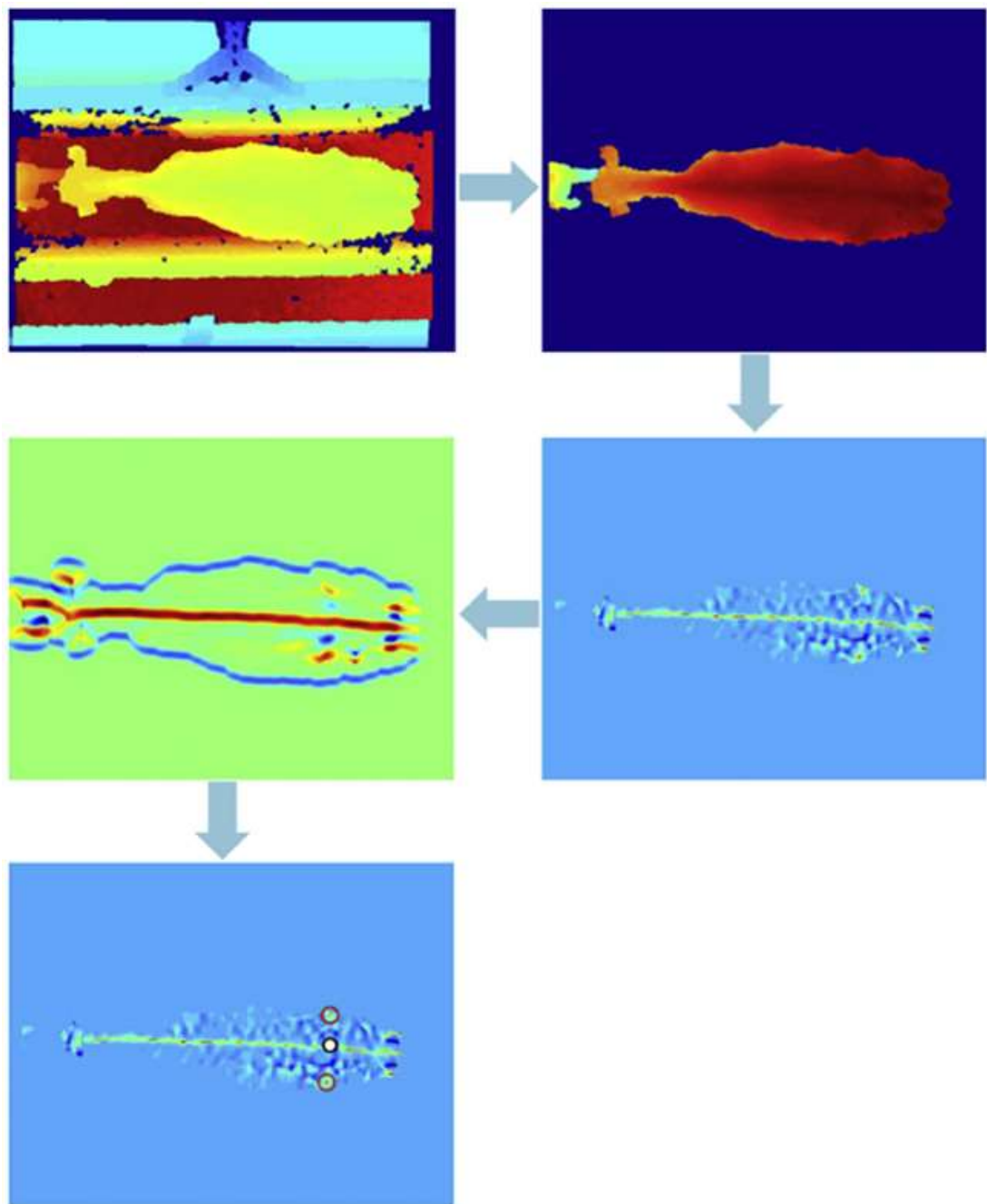


Abbildung 4: Darstellung der automatischen Berechnung der Wirbelsäule und der Hüfthöcker aus einem 3D Bild einer Kuh (JABBAR et al. 2017). Das erste Bild zeigt ein unbearbeitetes Tiefenbild, gefolgt von demselben Bild ohne Hintergrund. Darauf folgt die Krümmungsberechnung mit dargestellten Höhepunkten, welche anschließend in ein Binärbild umgewandelt werden und so die Wirbelsäule sichtbar machen. Am Ende steht ein Auswahlbild auf welchem die Krümmungen (höchste konvexe Regionen d.h. Wirbelsäule und Hüfthöcker) deutlich sichtbar sind.

Dynamische Messungen der Bewegung der Hüfthöcker und der Wirbelsäule lassen Rückschlüsse auf die Gleichmäßigkeit der Bewegung zu. Es wurden mittlere und

maximale Schwankungen (Höhenunterschiede) dieser Punkte berechnet. Dabei stellte sich die maximale Schwankung als sensiblere Messung für kleine Veränderungen, insbesondere bei Lahmheiten, heraus. Der Anteil der erfolgreich segmentierten Bilder lag bei 85,7 %. Die Kontrolle der Darstellung dieser Berechnungspunkte ermöglichte die Identifikation von Fehlerquellen und eine entsprechende Anpassung des Algorithmus. Dadurch verbesserte sich die Erkennungsrate der Hüfthöcker auf 96,1 % und der Wirbelsäule auf 100 %. Schlussendlich erreichte das System eine Sensitivität von 100 %, eine Spezifität von 75 % und eine Genauigkeit von 95,7 % bei einer Stichprobe von 23 Tieren.

VAN HERTEM et al. (2017) haben die Funktionsweise des von VIAZZI et al. (2014) und VAN HERTEM et al. (2014) entwickelten Systems zur Lahmheitsbeurteilung mittels 3D-Daten von September 2013 bis August 2014 in einem Milchviehstall in Belgien unter Praxisbedingungen getestet. Der 3D Sensor (Kinect™, Microsoft corp., Redmond WA, USA) wurde in einer Höhe von 3,45 m befestigt. Die Identifikation der Kühe erfolgte mittels einer RFID-Antenne, welche auch automatisch die Videoaufnahmen startete. Die Aufnahmen wurden mit 30 Bildern pro Sekunde gemacht. Alle Aufnahmen beinhalteten Tiefenbilder und wurden als oni-files gespeichert. Vor der Auswertung der Videos erfolgte eine automatische Filterung, wie von ROMANINI et al. (2013) beschrieben, um Bilder zu entfernen, auf denen a) mehrere Kühe zu sehen sind, b) die Kühe unregelmäßig laufen (stehen oder rennen) und c) nicht genug Bilder für die Auswertung vorhanden sind. Der gesamte Prozess - vom Starten der Videoaufnahmen, über die Einzeltiererkennung und Filterung der Videodaten bis hin zur Auswertung der Bilder mit dem Algorithmus - lief automatisch. Bei 62 % der Videos erfolgte eine automatische Berechnung der Variablen zur Ausgabe eines Locomotion Scores, 38 % wurden durch den Filter aussortiert. Zu 81 % erfolgte die Aussortierung aufgrund einer nicht möglichen Segmentierung des Kuhkörpers, weil die Kühe sich im Versuchsgang stauten. Durchschnittlich konnten in $49,3 \% \pm 10,8 \%$ aller Bilder die Variablen zur Lahmheitsbeurteilung berechnet werden. Im Mittel wurde für 79,4 % der Kühe mindestens fünfmal pro Woche automatisch ein Locomotion Score berechnet. Dabei stellte sich heraus, dass Fersen und Kühe von alten Müttern (Laktation ≥ 6) einen negativen Einfluss auf die Analyserate haben, wohingegen alte Kühe einen positiven (4. Laktation) oder geringfügigen (2., 3. und 5. Laktation) Einfluss haben. Es wurde festgehalten, dass Laktation, Laktationsstadium, Herdengröße und Melkdauer den größten Einfluss auf die Analyserate haben. Des Weiteren wurden die einzelnen

Merkmalsvariablen, die zur Berechnung des Locomotion Scores genutzt wurden, verglichen und es stellte sich heraus, dass der Krümmungswinkel des Rückens um die Hüftgelenke ($AUC = 0,719$) und die Rückenkrümmungsmessung des gesamten Rückens ($AUC = 0,702$) sich gut zur Klassifizierung eignen.

SCHLAGETER-TELLO et al. (2018) haben unter dem Versuchsaufbau von VAN HERTEM et al. (2017) weitere Untersuchungen durchgeführt. Es wurde getestet, ob das automatische System Kühe besser als lahm bzw. nicht lahm einstufen kann als ein menschlicher Beobachter. Dafür wurden die Ergebnisse von zwei Beobachtern miteinander verglichen sowie die Ergebnisse der Beobachter mit denen des automatischen Systems. Die Beobachter und das automatische 3D-System zeigten eine ähnliche Sensitivität für das Erkennen einer lahmen Kuh. Beobachter 1 und Beobachter 2 erreichten im direkten Vergleich eine Sensitivität von 74,2 %, die Beobachter im Vergleich mit dem automatischen System 73,9 – 71,8 %. Die Spezifität zwischen den Beobachtern lag bei 88,5 %, die Genauigkeit bei 84,2 %. Zwischen den Beobachtern und dem 3D-System lag die Spezifität bei 65,3 – 67,8 % und die Genauigkeit bei 67,7 – 69,2 %. In der allgemeinen Erkennung von lahmen Kühen scheint das automatische System demnach die Beobachtung durch den Menschen ersetzen zu können.

HANSEN et al. (2018) haben versucht mittels eines 3D Sensors drei Parameter zur gleichen Zeit zu messen und auszuwerten. Diese waren Lahmheit, Körperkondition und Gewicht. Dafür nutzten sie eine 3D Kinect-like Tiefenbild Kamera, die aus 2,3 m Höhe 30 Bilder pro Sekunde aufnahm, welche 2,5 m Ganglänge zeigten. Die Technik war in einem IP66 wassergeschützten Gehäuse angeordnet. Eine Tiererkennung über RFID startete automatisch die Aufnahmen. Bevor die Bilddaten ausgewertet werden konnten, mussten diese vorbereitet werden. Das heißt, der reine Kuhkörper musste aus dem Bild extrahiert werden. Anschließend erfolgte ein Vorlagenvergleich, um herauszufinden, ob auf dem Bild das Vorderteil, das Hinterteil oder beides zusehen ist. Dabei zeigte sich, dass Krümmungsdaten sich besser zum Vorlagenvergleich eignen als Tiefenbilddaten (Abbildung 5). Vor der weiteren Verarbeitung des Bildes wurde das Bild so gedreht, dass die Wirbelsäule horizontal im Bild dargestellt wird.

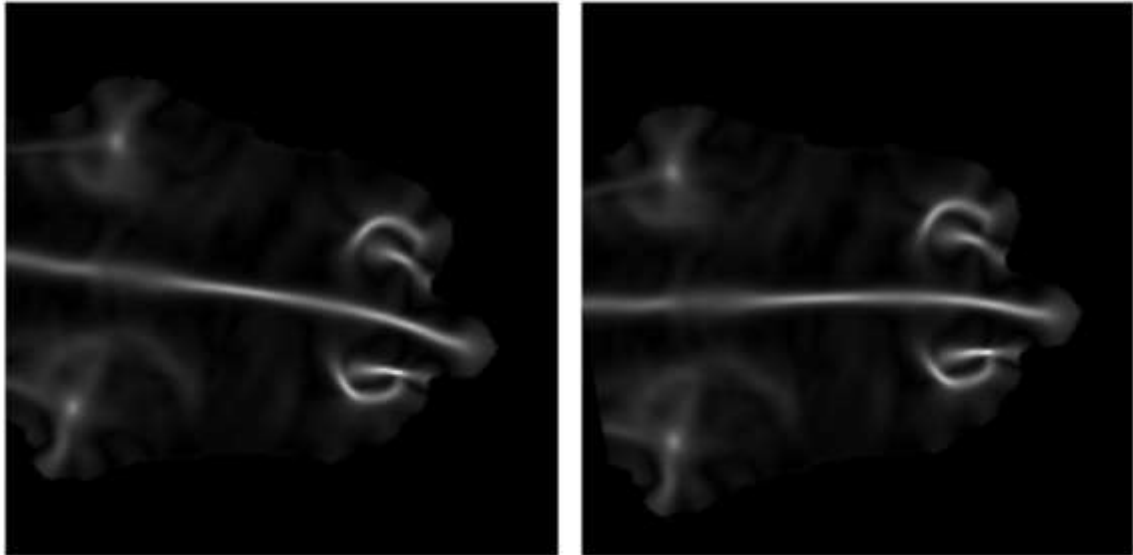


Abbildung 5: Krümmungsdaten berechnet aus einem Tiefenbild einer Kuh (HANSEN et al. 2018). Vor (links) und nach (rechts) der automatischen Ausrichtung des Bildes. Die am stärksten gekrümmten Bereiche werden deutlich gezeigt.

Zur Lahmheitsbeurteilung wurde die Wirbelsäule extrahiert und die Krümmungsinformationen genutzt, um ein Polynom zweiten Grades daran anzupassen. Aus diesen Daten ließ sich die Schwere der Lahmheit abschätzen und - falls erwünscht - ein korrelierter Mobilitätsscore. Die Krümmung konnte sowohl in der vertikalen x-z-Ebene als auch in der horizontalen x-y-Ebene generiert werden, sodass eine schlangenähnliche 3D Kurve entsteht, die sich durch den Raum bewegt, wenn sich die Kuh vorwärtsbewegt. Der Schweregrad der Lahmheit war abhängig vom Ausmaß der Krümmung. Darüber hinaus soll die Kurvenform dazu genutzt werden können, um die lahme Gliedmaße zu lokalisieren, woran aktuell aber noch geforscht wird. Schlussendlich konnte eine Genauigkeit von 83 % erreicht werden (wenn ≥ 3 = lahm).

2.5.5 Validierung automatischer Locomotion Scores

Die Validierung von automatischen Scores erfolgt anhand der Gegenüberstellung von automatisch berechneten Scores und visuell erfassten Scores. Die Vorgehensweise bei der Validierung und visuellen Bonitur ist nicht immer gleich. Im nachfolgenden wird die Vorgehensweise in ausgewählten Studien beschrieben.

Bei RAJKONDAWAR et al. (2006) wurde durch einen Veterinär einmal pro Woche eine Gangbeurteilung sowie eine Beurteilung der Klauen vorgenommen. Läsionen wurden, wenn notwendig, behandelt und die Gangbeurteilung erfolgte für jedes Bein einzeln. Kühe, die an zwei aufeinanderfolgenden Wochen die Note 1 für alle vier Beine

erhielten, wurden aus der Versuchsgruppe ausgeschlossen. Für die Gangbeurteilung wurden die Methoden von SPRECHER et al. (1997) und WELLS et al. (1993) eigens modifiziert. Da nur wenig Scores > 3 vorkamen, wurden Scores mit 4 und 5 aus der Studie ausgeschlossen.

Bei PASTELL UND KUJALA (2007) wurden 10.000 Daten von 73 Kühen über fünf Monate erfasst. Zur visuellen Kontrolle der automatischen Daten wurde jede Woche ein Locomotion Scoring nach SPRECHER et al. (1997) durchgeführt und lahme Kühe auf Läsionen an der Klaue untersucht. Kühe mit einem Score von 1 oder 2 wurden als gesund angesehen und Kühe mit einem Score von 3 oder mehr als lahm.

PASTELL et al. (2008) hatten eine Stichprobe mit 43 Kühen, deren Bewegung von einem Veterinär nach SPRECHER et al. (1997) beurteilt wurde.

Bei BICALHO et al. (2007) wurde das visuelle Locomotion Scoring in Studie 1 durch einen Veterinär anhand einer 5-Punkte Bewertung durchgeführt, in Studie 2 von drei Veterinären. Die Kühe wurden beim Verlassen des Melkstandes beobachtet und das Ergebnis in ein Diktiergerät gesprochen. Zur Ermittlung des visuellen Scores von den drei Beobachtern, wurde der Mittelwert gebildet und anschließend auf eine ganze Zahl gerundet.

BRENNINKMEYER et al. (2007) haben die Zuverlässigkeit der visuellen Bewertung des Locomotion Scorings untersucht. Dabei wurden vier Beobachter, drei davon unerfahren, darauf trainiert das Locomotion Scoring nach einer modifizierten Beschreibung von WINCKLER UND WILLEN (2001) durchzuführen. Diese gibt die Möglichkeit der Einstufung in fünf Kategorien. Die Ergebnisse zeigten, dass die Übereinstimmung zwischen den Beobachtern mit der Anzahl der Durchführungen steigt. Die beste Übereinstimmung wurde nach dem letzten Durchlauf mit der meisten Übung erreicht. Wird die Skala der Bewertungsmöglichkeiten auf vier oder zwei minimiert, zeigten bereits die Ergebnisse der relativ untrainierten Beobachter eine zufriedenstellende Übereinstimmung.

BAHR et al. (2008) haben drei erfahrene Beobachter eingesetzt, um das visuelle Locomotion Scoring anhand der aufgenommenen Videos durchzuführen. Es wurde nach der Methode von WINCKLER UND WILLEN (2001) vorgegangen, jedoch wurden statt der fünf Stufen nur drei vergeben (1 = gesund, 2 = mittelmäßig lahm, 3 = schwer lahm). Jeder Beobachter hat jede Kuh zweimal bewertet und der Durchschnitt dieser sechs Bewertungen pro Kuh wurde für die Auswertung genutzt.

Bei SONG et al. (2008) wurde die visuelle Evaluation von vier Beobachtern vorgenommen, welche die Kühe bewerteten, während sie durch die Versuchsumgebung liefen. Zusätzlich haben zwei der Beobachter die Kühe noch einmal anhand der Videoaufnahmen bewertet.

ITO et al. (2010) haben sich beim visuellen Locomotion Scoring auf drei Merkmale beschränkt. Diese waren ein asymmetrischer Gang, Kopfnicken und die Verlagerung des Gewichts auf bestimmte Beine. In der Auswertung wurden zwei neue Variablen gebildet. Bei der einen wurden Kühe mit einem Score ≥ 3 als lahm bezeichnet und bei der anderen Kühe mit einem Score ≥ 4 .

Bei POURSAHERI et al. (2010) wurden alle in der Literatur als Lahmheitsindikator genannten Merkmale betrachtet. Diese waren: Druckempfindlichkeit, Rückenkrümmung, verlangsamte Bewegung, ungleichmäßiger Gang, verringerter Übertritt, erhöhte Abduktion und Kopfnicken. Sobald eine Kuh einen dieser Indikatoren zeigte, wurde der Score 2 gegeben. Beim Auftreten mehrerer Indikatoren oder eines Indikators in sehr starker Ausprägung wurde der Score 3 gegeben. Nur wenn keiner dieser Indikatoren in der Bewegung sichtbar war, wurde eine 1 vergeben.

Bei PLUK et al. (2012) wurden die aufgenommenen Videos zweimal pro Woche von einem Beobachter ausgewertet und Locomotion Scores von 1 bis 3 vergeben.

Die Validierung bei VIAZZI et al. (2014) wurde von einem gut ausgebildetem Veterinär anhand der 2D Videoaufnahmen durchgeführt. Dabei wurde ein Scoring mit den Noten 1 – 5 genutzt und die Rückenkrümmung, die Kopfhaltung, das Übertreten, die Gewichtsverteilung auf die Beine und allgemein die Asymmetrie des Ganges beurteilt. Kühe mit einem Score von 1 und 2 wurden als „nicht lahm“ klassifiziert und Kühe mit einem Score > 3 als „lahm“.

VAN HERTEM et al. (2014) haben das Locomotion Scoring nach SPRECHER et al. (1997) angewendet und die Kühe parallel zu den Videoaufnahmen live beurteilt. Zwei erfahrene Beobachter haben alle Kühe der Herde in wöchentlichen Intervallen bewertet. Insgesamt haben 41 Beurteilungstage stattgefunden.

Auch bei JABBAR et al. (2017) erfolgte das visuelle Locomotion Scoring nach der Methode von SPRECHER et al. (1997). Die Bewertung erfolgte in zwei Schritten. Zum einen wurde eine live Bewertung im Laufstall vorgenommen, kurz nachdem die Kühe unter dem Sensor entlanggelaufen sind, zum anderen haben zwei 2D Kameras die Kühe

parallel zum 3D Sensor, mit Blickwinkel von der Seite und auf den Rücken der Kuh, aufgenommen, um den finalen, visuellen Locomotion Score zu vergeben und die Videos den Kühen anhand der Brandnummern zuordnen zu können.

In der Studie von VAN HERTEM et al. (2017) bewerteten zwei geschulte Beobachter abwechselnd die Kühe wöchentlich nach der Methode von SPRECHER et al. (1997). Da zur Modellentwicklung und Auswertung zweistufige Daten notwendig sind, wurden Tiere mit einem Locomotion Score von 1 und 2 als nicht lahme Kühe zusammengefasst und jene mit einem Score > 3 als lahme Kühe.

Die Untersuchungen von SCHLAGETER-TELLO et al. (2018) erfolgten auf der Grundlage von zwei Versuchsdurchgängen mit je zwei aufeinanderfolgenden Versuchstagen. Es wurde jeweils von allen Kühen nach der morgentlichen Melkung ein automatischer und ein manueller Locomotion Score vergeben, sowie alle vier Klauen begutachtet und geschnitten. Insgesamt lagen Daten von 223 Kühen vor. Das Locomotion Scoring wurde von zwei wissenschaftlichen Beobachtern live etwa zehn Meter hinter dem 3D Sensor nach der Methode von FLOWER UND WEARY (2006) durchgeführt. Es wurden Scores von 1 – 5 vergeben, wobei ein Score ≥ 3 als lahm klassifiziert wurde.

HANSEN et al. (2018) haben ihre Versuche an 200 Holstein-Friesian Kühen durchgeführt, die einmal unter dem Sensor entlanggelaufen sind. Drei trainierte Mitarbeiter haben das visuelle Locomotion Scoring nach SPRECHER et al. (1997) direkt nach dem Melken durchgeführt. Um die Wiederholbarkeit der Ergebnisse zu testen, wurden 16 Kühe fünfmal hintereinander an einem Tag unter dem System entlang geleitet. Jeder Testlauf hat weniger als eine Stunde gedauert und es konnten nicht für alle Kühe automatische Scores generiert werden. Überwiegend wurde der Misserfolg durch das Verhalten der Kühe verursacht, beispielsweise wenn eine Kuh den Kopf auf den Rücken der anderen gelegt hat, aber auch, wenn das RFID-System eine Kuh nicht erkannt hat. Das Ziel war die Sicherstellung der Erkennung fehlerhafter Daten und der automatische Ausschluss dieser aus der Analyse.

3. Material und Methode

3.1 Tiermaterial

Die Untersuchungen wurden auf einem Milchviehbetrieb in Brandenburg in der Nähe von Nauen durchgeführt. In der Zeit von Juli bis Dezember 2016 umfasste die Herde durchschnittlich 349 Kühe der Rasse Deutsche Holstein (schwarzbunt), wovon sich 304 in Laktation befanden. Das Leistungsniveau der Herde lag im Durchschnitt bei 25,4 kg Milch je Kuh und Tag mit einem Fettgehalt von 3,75 % und einem Eiweißgehalt von 3,26 %.

3.2 Versuchsumgebung

Aufgrund der Gegebenheiten vor Ort wurde für die Versuchsdurchführung mit den Video- und Sensordaten der Treibgang, welcher nach dem Melken von den Kühen durchlaufen wird, ausgewählt (Abbildung 6). Der Gang hat eine Breite von 1,28 m, eine Länge von 17,30 m und eine Deckenhöhe von 2,40 m. Nach 3,48 m befindet sich auf der linken Seite eine 1,62 m breite Öffnung. Diese kann durch ein einfaches Metalltor, welches lediglich aus zwei Rohren besteht und somit keine visuelle Barriere darstellt, verschlossen werden. Die Kuhgruppen 1 und 3 laufen täglich an dieser Öffnung vorbei. Alle anderen Gruppen gehen entweder durch diese Öffnung zurück in ihre Gruppen oder verlassen den Melkstand auf der anderen Seite. Einmal pro Woche durchlaufen alle melkenden Kühe diesen Gang komplett, da im hinteren Teil das Klauenbad platziert wird.

Da im diesem Gang einige Elemente des Spaltenbodens wackelten, wurden diese vor Versuchsbeginn ausgetauscht, um die Tiere nicht zu beeinflussen und einen rutschfesten, ebenen Untergrund zu bieten, wie es in der Literatur empfohlen wird.

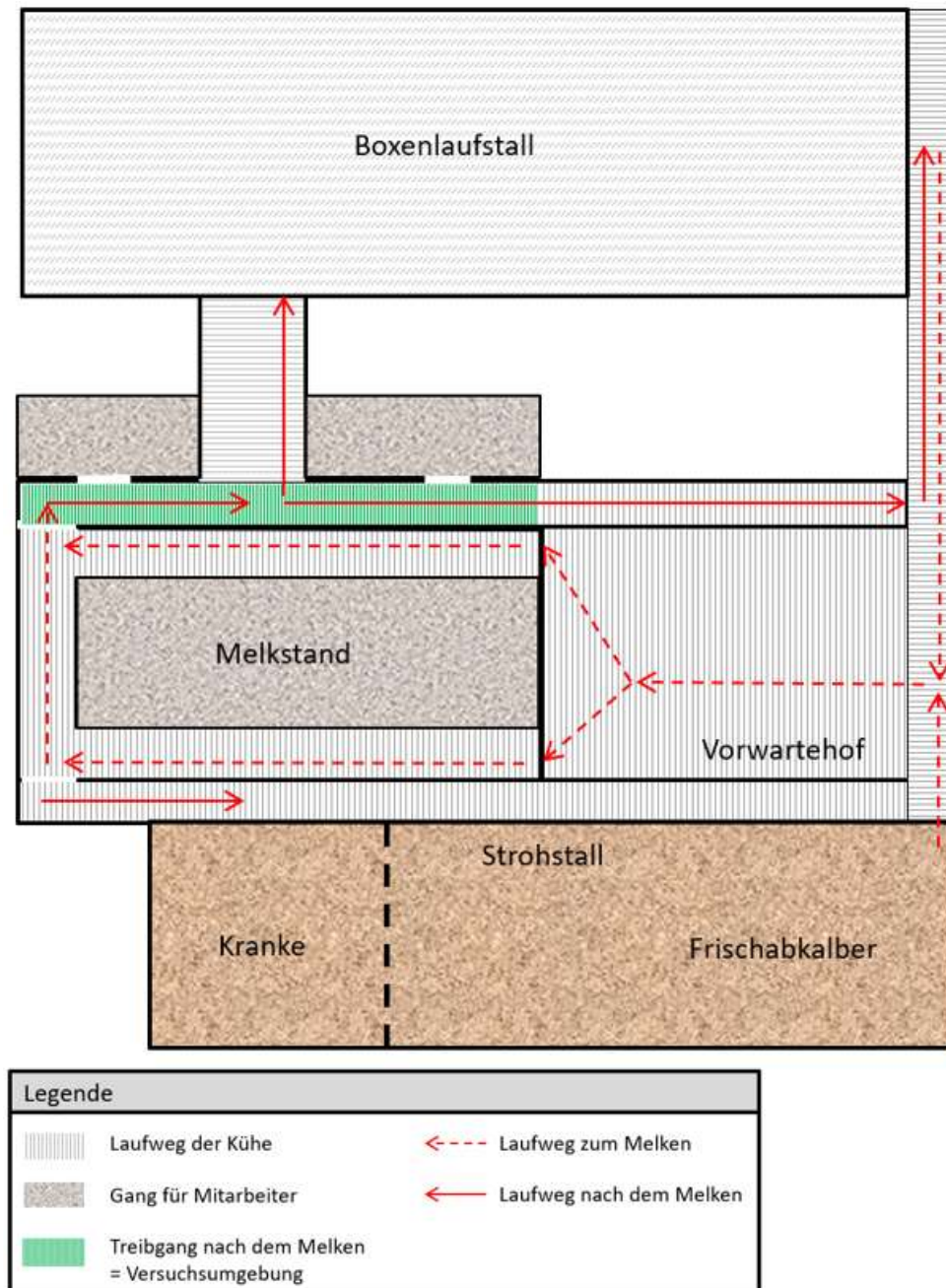


Abbildung 6: Skizze der Versuchsumgebung mit Darstellung der Laufwege der Kühe

3.2.1 Haltung

Die Haltung der Milchkühe erfolgt in einem alten LPG Stall. Die Hochleistungskühe stehen im Boxenlaufstall mit Außenfütterung, die Lahmen und Frischmelker haben einen separaten Bereich auf Tiefstreu, welcher durch ein Absperrgitter voneinander getrennt ist. Im Boxenlaufstall sind Hochliegeboxen mit Gummimatten integriert. Die Liegeboxen werden mit Häckselstroh und Kalk eingestreut. Die Laufgänge bestehen aus Gussasphalt und werden durch einen Faltschieber entmistet. An den Stirnseiten des

Stalls sind Windschutznetze angebracht, um Durchzug zu vermeiden. Im Sommer werden Ventilatoren eingeschaltet, um eine höhere Lufrate zu gewährleisten und den Kuhkomfort zu erhöhen. Für die Wasserversorgung sind in jeder Gruppe mindestens zwei von Schwimmern gesteuerte Kipptränken vorhanden. Das Tier:Liegeplatz-Verhältnis ist 1 : 1,1.

Die Kühe sind in sechs Gruppen aufgeteilt. Die Frischabkalber, Lahmen und Euterkranken stehen im Strohstall, um einen möglichst kurzen Weg zum Melkstand zu gewährleisten. Die gesamte Strecke vom Strohstall zum Melkstand und zurück beträgt ca. 80 m. Der Boxenlaufstall ist in vier Gruppen mit drei verschiedenen Futterrationen unterteilt. Die Jungkühe und Gruppe 1 bekommen Hochleistungsfutter, Gruppe 2 Mittelleistungsfutter und Gruppe 3 Futter für Altmelker. Gruppe 1 - 3 und die Jungkühe müssen zum Melkstand und zurück etwa eine Strecke von 200 m gehen. Einmal pro Woche verlängert sich der Weg für Gruppe 2 und die Jungkühe um etwa 75 m, da auf dem Rücktriebweg durch ein Klauenbad gelaufen wird.

Alle Kühe werden zweimal täglich in einem 2x12 Fischgrätenmelkstand der Firma SAC gemolken. Das Melken erfolgt jeweils von 02:30 bis 07:30 Uhr und 14:00 bis 18:00 Uhr.

3.2.2 Fütterung

Die Fütterung basiert auf einer totalen Mischration (TMR), welche zweimal täglich mittels eines Futtermischwagens vorgelegt wird.

Der Grobfutteranteil in der TMR setzt sich aus Maissilage und Grassilage zusammen. Als Saftfutter beinhaltet die TMR Feuchtmais und Pressschnitzel. Das Kraftfutter besteht aus Rapsschrot, Getreideschrot und Sojaschrot. Außerdem werden Futterkalk und Futterfett dazu gegeben.

3.2.3 Klauenmanagement

Der Klauenkorrekturschnitt erfolgt als Bestandsschnitt zweimal pro Jahr durch einen professionellen Klauenpfleger und die Befunde werden digital überlassen. Zwischen diesen Terminen werden durchschnittlich einmal pro Woche lahme Tiere selbst vom Betriebsleiter im eigenen Behandlungsstand ausgeschnitten. Treten vermehrt lahme

Tiere auf, wird der Klauenpfleger zwischen den Bestandsschnitten zusätzlich auf den Betrieb bestellt.

Zur Pflege der Klauen wird im Betrieb einmal pro Woche ein Klauenbad durchgeführt. Derzeit werden vier Durchlaufwannen genutzt. Eine große Wanne mit Wasser zum Vorsäubern der Klauen und im Anschluss drei kleinere Wannen mit dem Behandlungsmittel. Die erste Wanne mit dem Reinigungswasser wird nach Durchlauf der Hälfte der Tiere einmal gewechselt. Während des Versuchszeitraums wurde das Mittel zur Behandlung der Klauen einmal gewechselt. Von August bis Oktober wurde MS Formades mit dem Wirkstoff Formaldehyd verwendet, von Oktober bis Dezember das Mittel Podocur Extra mit den Wirkstoffen Glutaral und 2-Phenylphenol. Das Klauenbad wird im Treibgang, der nach dem Melkstand durchlaufen wird, aufgestellt.

3.3 Technik

3.3.1 3D Sensor

Zur automatischen Berechnung des Locomotion Scores wurden zwei 3D/4D-Sensoren der Marke ASUS xTion Pro Live verwendet (Abbildung 7). Der Sensor funktioniert nach dem Prinzip des strukturierten Lichts, wobei ein vordefiniertes Muster von Infrarot-Punkten von einem eingebauten Projektor projiziert wird. Dieses Muster wird von einer Infrarot-Kamera erfasst. Daraus wird die Disparität jedes Punktes berechnet und die Tiefe hieraus abgeleitet.



Abbildung 7: Asus xTion Pro Live

Eine xTion liefert 2,5D-Tiefenbilder mit einer Auflösung von 640 x 480 Pixeln mit einer Abtastungsrate von 30 Hz und hat eine Tiefengenauigkeit von 5 – 15 mm bei einem Messabstand von 1 m (GONZALEZ et al. 2013). Jeder Pixel hat einen 16-bit Wert, der dem Abstand in Millimeter zwischen dem 3D Sensor und dem entsprechenden Punkt auf dem Objekt entspricht.

3.3.2 Tiererkennung

Die Tiererkennung wurde mittels radio-frequency identification (RFID) sichergestellt. Dafür wurden die Kühe zufällig über alle Gruppen am Halsband mit den RFID-Sensoren ausgestattet (Abbildung 9). Insgesamt wurden 386 Tiere markiert. Die Identifizierung der Tiere erfolgte über elektromagnetische Wellen. Wenn eine Kuh mit einem RFID-Chip eine Antenne passierte, registrierte diese automatisch die Nummer sowie Datum und Uhrzeit und speicherte diese ab. Durch die Montage von zwei Antennen in der Versuchsumgebung, wird der Treibgang von 1,28 m auf 0,95 m verschmälert (Abbildung 8).



Abbildung 8: Antennen der Tiererkennung im Treibgang



Abbildung 9: RFID Sensor am Halsband einer Kuh

3.3.3 2D Kameras

Es wurden zwei 2D-Industrie-Kameras in die Messumgebung integriert, welche Bilder mit einer Auflösung von 1280 x 1024 Grauwert-Pixeln liefern (Abbildung 10).



Abbildung 10: Im Stall montierte 2D Kamera

3.4 Visuelles Locomotion Scoring

Um die automatisch berechneten Scores zu validieren, wurde diesen eine visuelle Bonitur als Goldstandard gegenübergestellt. Dabei wurde die Bewertungsmethode nach SPRECHER et al. (1997) (siehe Tabelle 1) angewandt. Die Kühe wurden jedoch nicht bewertet, wenn sie standen, sondern ausschließlich in der Bewegung. Das wichtigste Merkmal für die Bonitur war der Rücken der Kuh. Beschrieb dieser eine gerade Linie, wurde die Note 1 vergeben, war er leicht gekrümmt, wurde schon die Note 2 vergeben. War der Rücken gekrümmt und der Gang durch verkürzte Schritte verändert, wurde die Note 3 vergeben. Bei einem stark gekrümmten Rücken und der Verzögerung des Ganges durch die Bevorzugung eines Beines wurde der Gang mit einer 4 bewertet. Konnte die Kuh kaum laufen, wölbte den Rücken stark und hat ein Bein kaum belastet, musste mit einer 5 bewertet werden.

Alle in dieser Studie zur Auswertung stehenden Bonituren wurden von einer wissenschaftlichen Mitarbeiterin durchgeführt, die bereits langjährige Erfahrungen im Locomotion Scorings hat und vor dem Start der Versuche noch einmal gesondert geschult wurde. Eine kleine Stichprobe (Stichprobe 6) wurde zusätzlich von zwei weiteren wissenschaftlichen Mitarbeitern - einer erfahren, der andere unerfahren - bonitiert.

3.5 Statistische Auswertung

Die Daten der automatischen und der visuellen Locomotion Scores wurden in Microsoft Excel[®] (2010 Microsoft Corporation) aufbereitet. Die erste Auswertung dieser Daten

erfolgte mit Excel[®] und die statistische Auswertung der Daten mit der Statistik-Software R (R Core Team, 2015).

Mit R wurde die Korrelation nach Spearman zwischen den automatischen und den visuellen Locomotion Scores berechnet, weil dieser Rangkorrelationskoeffizient robust gegenüber Ausreißern ist.

Außerdem wurde eine Varianzanalyse vorgenommen, um die Unterschiede der Messwerte zwischen den einzelnen Scores zu ermitteln. Der Faktor in der Varianzanalyse der vorliegenden Arbeit ist der visuelle Locomotion Score und die Zielvariable der Messwert der Rückenkrümmung. Die Nullhypothese besagt, dass kein Unterschied zwischen den Messwerten der Rückenkrümmung in den Gruppen der visuellen Scores besteht.

Zur Bestimmung von signifikanten Unterschieden zwischen den Gruppen, wurde der Tukey-Test als post-hoc Test der Varianzanalyse durchgeführt. Ist der p-Wert $< 0,05$, ist der Unterschied zwischen den untersuchten Gruppen statistisch signifikant.

4. Ergebnisse

Prozessentwicklung

4.1 Entwicklung des experimentellen Designs

In der **ersten Versuchsphase** wurde für die Position des 3D Sensors eine Stelle kurz vor dem Ende der Versuchsumgebung gewählt. Die Position des Sensors befand sich dabei vor einer Wandöffnung, die der Beobachterin auf dem dahinter liegenden Mitarbeitergang. Anschließend wurden zwei weitere Positionen durch die Beobachterin getestet. Zum einen im großen Durchgang mit größerem Abstand zur Versuchsumgebung kurz vor dem Boxenlaufstall und zum anderen auf dem Mitarbeitergang vor dem großen Durchbruch in der Versuchsumgebung (Abbildung 11).

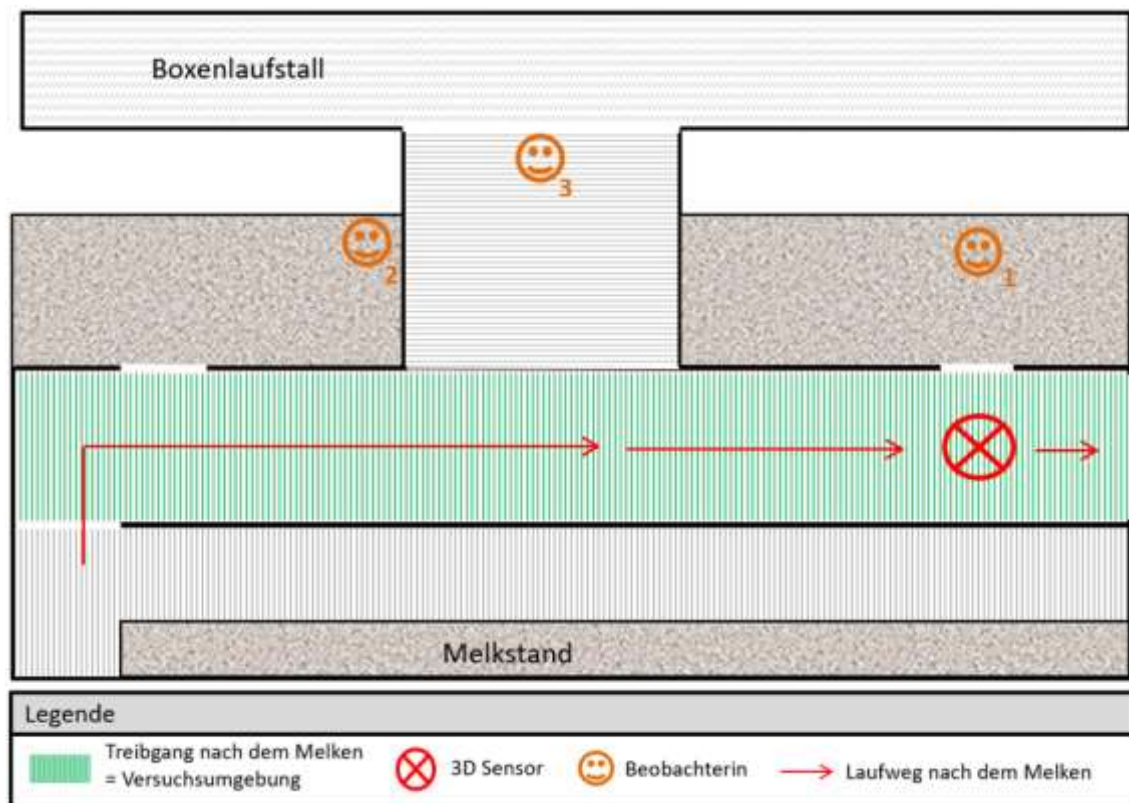


Abbildung 11: Skizze der Versuchsumgebung in der ersten Versuchsphase. Darstellung der Position des 3D Sensors und der Beobachterin im ersten Versuchszyklus

Die dritte Position der Beobachterin wies dabei die beste Perspektive zur Durchführung des visuellen Locomotion Scorings auf.

In der **zweiten Phase** wurde die Versuchsanordnung geändert und die Position des 3D Sensors auf Höhe des großen Durchgangs verlegt (Abbildung 12).

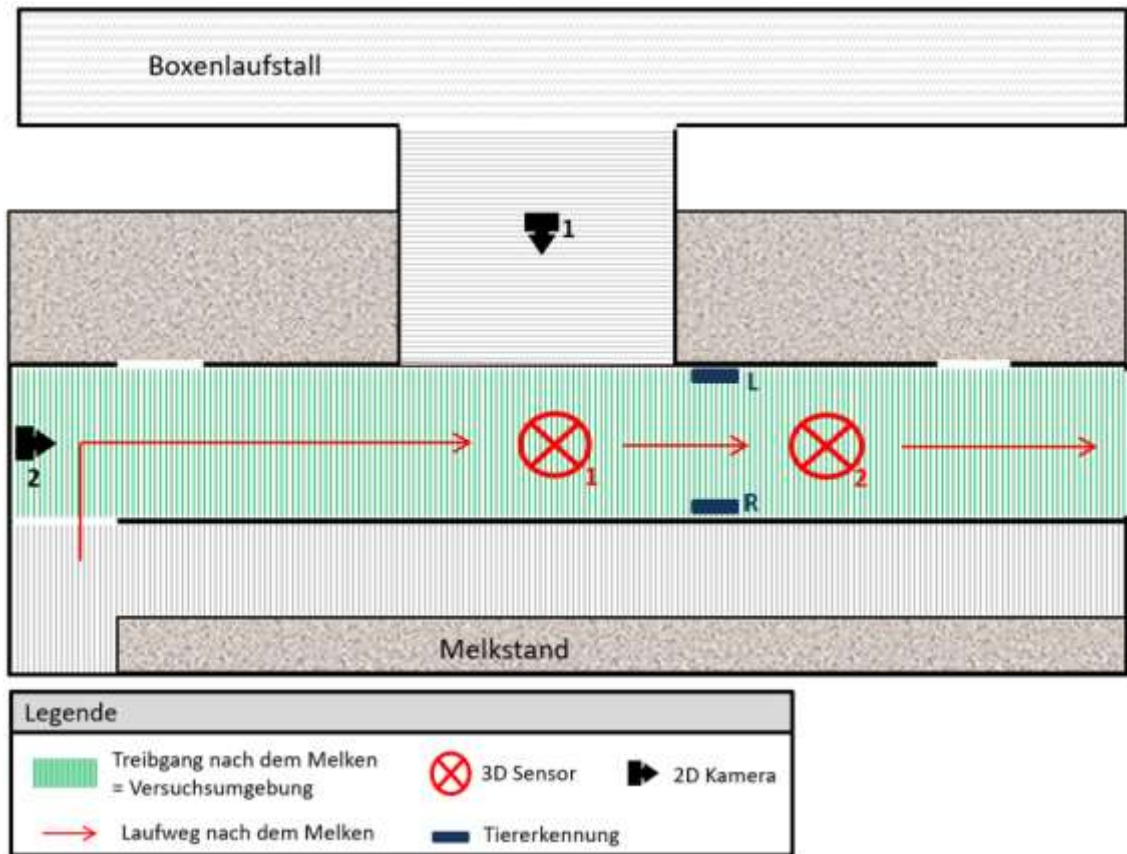


Abbildung 12: Skizze der Versuchsumgebung in der zweiten Versuchsphase. Darstellung der Position der 3D Sensoren, 2D Kameras und der Tiererkennung

Zusätzlich wurden zwei 2D Kameras montiert. Beide Kameras wurden in einer Höhe von 2,2 m montiert. Kamera 1 war 2,5 m vom Versuchsgang entfernt, Kamera 2 hat den gesamten Versuchsgang visualisiert.

In der neuen Versuchsanordnung wurden zwei 3D Sensoren hintereinandergeschaltet. Sie wurden so positioniert, dass die Kuhrücken jeweils komplett erfasst werden können, sich die beiden Messbereiche aber nicht überschneiden. Die 3D Sensoren wurde in einer Höhe von 2,80 m in einem Schutzgehäuse über dem Versuchsgang angebracht (Abbildung 13).



Abbildung 13: Schutzgehäuse des 3D Sensors - Aufnahme oberhalb der Decke am Versuchsende

Das Schutzgehäuse wurde aus Aluminium gefertigt. Im Schutzgehäuse befand sich eine Isolier-Fensterglasscheibe, auf welcher der 3D Sensor positioniert wurde (Abbildung 14).



Abbildung 14: Schutzgehäuse des 3D Sensors von unten durch die Decke aufgenommen. Die Aussparung im Aluminiumgehäuse ist mit einer Isolier-Fensterglasscheibe abgedeckt. Darüber ist der 3D Sensor.

Zur Automatisierung der Prozessabläufe und besseren Datenverarbeitung wurden in das Schutzgehäuse neben dem 3D Sensor auch ein Rechner (Shuttle DS Barebone Systems) und ein Netzwerk-Switch (NETGEAR ProSAFE) platziert (Abbildung 15).

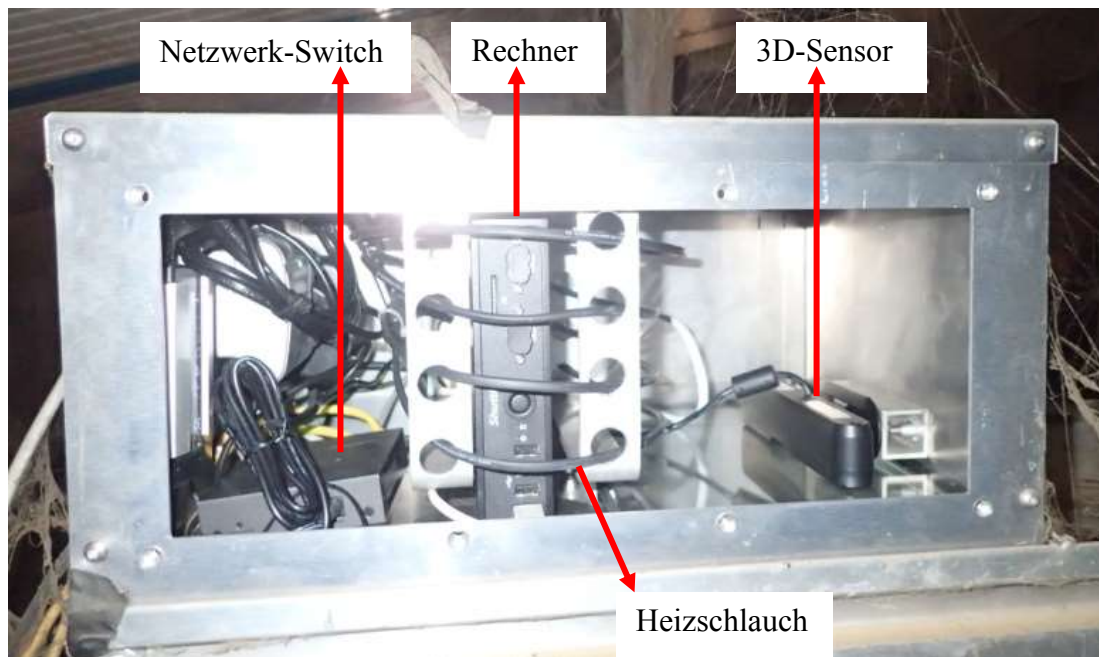


Abbildung 15: Darstellung der Anordnung der Messtechnik im Schutzgehäuse

Da das gesamte technische Equipment zur reibungslosen Funktion in den jeweiligen Spezifikationen Anforderungen an den Temperaturbereich stellt, wurde zur Vermeidung von Temperaturen unterhalb des Gefrierpunkts ein Heizschlauch um den Rechner gelegt (Abbildung 15), welcher sich bei $< +5\text{ }^{\circ}\text{C}$ einschaltet und zur Vermeidung von Temperaturen über $35\text{ }^{\circ}\text{C}$ Wärmtauscher auf dem Schutzgehäuse montiert.

Statt der visuellen Beobachterin wurden zwei 2D Kameras montiert, welche das Geschehen in der Versuchsumgebung aufgezeichnet haben. Die Kameras wurden so positioniert, dass die Aufnahmen exakt den Moment wiedergaben, in dem eine Kuh unter 3D Sensor 1 entlangging (Abbildung 12). Kamera 1 wurde in einer Höhe von 2,20 m am Anfang des Rücktriebgangs/der Versuchsumgebung montiert und zeichnete das Geschehen in der gesamten Versuchsumgebung auf. Kamera 2 wurde in der Mitte der großen Öffnung des Treibganges mit einem Abstand von zwei Metern in einer Höhe von 2,20 m an der Wand befestigt. Die plattenförmigen Antennen der Tiererkennung wurden zwischen den beiden 3D Sensoren an der Wand montiert. Wenn von den Antennen eine ID erkannt wurde, erfolgte die Weiterleitung der ID an das Messsystem und die Aufnahme aller 2D Kameras und 3D Sensoren wurde automatisch gestartet.

An dem Rechner wurden der 3D Sensor, die Tiererkennung und die 2D Kameras angeschlossen. Die Kameras und der 3D Sensor nahmen ständig Bilder auf, welche kontinuierlich im Ring-Buffer gespeichert wurden. Die Größe der einzelnen Ring-Buffer war auf 60 Bilder begrenzt. Sowohl die Kameras als auch der 3D Sensor arbeiteten mit einer Geschwindigkeit von 30 Bildern pro Sekunde. Daraus ergab sich eine Bildsequenz von zwei Sekunden pro Kamera/Sensor. Es wurde eine Software programmiert, welche die Speicherung der Bilddaten aus dem Ring-Buffer auf der Festplatte veranlasste, sobald eine RFID erkannt wurde. Die Speicherung der Bilddaten erfolgte auf einem separaten PC in Ordnern. Für jede RFID wurde ein Ordner angelegt (Abbildung 16). In diesen Kuh-Ordnern befanden sich weitere Ordner, welche mit dem Datum bezeichnet wurden, an welchem die darin enthaltenen Daten aufgezeichnet wurden. In diesen Ordnern befanden sich Unterordner, welche mit der Uhrzeit bezeichnet waren. Es konnten sich maximal zwei Uhrzeit-Ordner in einem Datum-Ordner befinden, da die Tiere zweimal am Tag nach dem Melken an der Tiererkennung vorbeikamen.

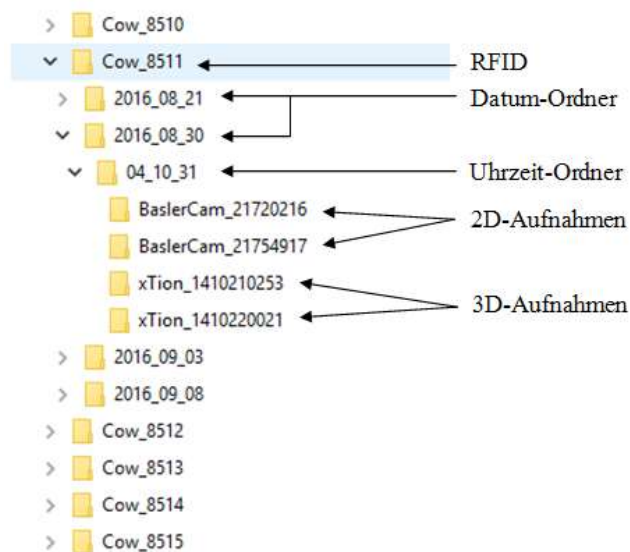


Abbildung 16: Darstellung der Ordnerstruktur für die Datenerfassung

In jedem Uhrzeit-Ordner befanden sich wiederum vier weitere Ordner. Die ersten beiden Ordner (BaslerCam_) enthielten die Bilddaten der beiden 2D Kameras und die letzten Ordner (xTion_) die des 3D Sensors.

Vom 03.08.2016 bis zum 30.12.2016 hat das System automatisch Videodaten nach diesem Prinzip gespeichert, welche die Datengrundlage der vorliegenden Dissertation darstellen.

Das **Versuchsdesign der abschließenden Validierungsversuche** entspricht dem von Abbildung 12, allerdings war nur die rechte Antenne der Tiererkennung aktiv. Diese hat alle RFID-Sender innerhalb vom maximal 40 cm Abstand erkannt und zusammen mit dem Datum und der Uhrzeit der Erkennung abgespeichert. Zur finalen Auswertung wurden nur die Daten von 3D Sensor 1 herangezogen.

4.2 2,5D- und 3D-Datenverarbeitung

Nach der Erfassung wurden die aufgenommenen 2,5D-Tiefenbilder mithilfe verschiedener Algorithmen weiterverarbeitet und die Körpermerkmale zur Ermittlung der Locomotion Scores abgeleitet. Die Algorithmen wurden vorrangig im 2,5D-Raum entwickelt, um die Auswertezeit und den Rechenaufwand zu minimieren. Anschließend wurden markante Körpermerkmale in den 3D-Raum überführt. Dadurch wurden die zeitaufwendigeren 3D-Algorithmen nur in sehr kleinen Datenbereichen angewandt.

Da nur der Kuhrücken für die Bewertung relevant war, wurde dieser zunächst komplett von der Umgebung separiert. Zu der Umgebung gehörten die Wände, der Boden und auch Körperteile von anderen Kühen. Mittels Berechnung eines Differenzbildes aus einem statischen Referenz-Tiefenbild ohne Kuh und einem Tiefenbild mit Kuh konnten der Boden und die Wände im 2,5D-Tiefenbild herausgefiltert werden. Da sich der Rücken einer Kuh in 1 – 2 m Höhe befindet, konnten die Punkte, deren Tiefe außerhalb dieses Tiefenbereichs lagen, ebenfalls gefiltert werden. Alle ausgeblendeten Pixel des Tiefenbildes wurden in einem Maskenbild zusammengeführt, welches im nächsten Schritt auf das Tiefenbild mit Kuh angewendet wurde. Hier waren außer der zur Kuh gehörenden Punkte auch noch zusätzliche Artefakte zu erkennen, welche aufgrund der niedrigen Präzision des 3D-Scanners noch nicht herausgefiltert werden konnten. Dieses Tiefenbild wurde anschließend in ein Binärbild umgewandelt, dessen Pixel nur die Farben Schwarz und Weiß annehmen. Die Pixel ohne Tiefeninformation bekamen die Farbe Weiß, Pixel mit Tiefeninformation die Farbe Schwarz. Auf dem Binärbild wurden Punkt-Cluster (zusammenhängende schwarze Bildbereiche) identifiziert. Wenn der Rücken einer Kuh in dem Bild enthalten war, bildete dieser das Cluster mit der größten Anzahl von Punkten. Alle übrigen Cluster konnten dann verworfen werden. Diese Methode funktionierte auch, wenn Körperteile anderer Kühe, die kleiner als der Rücken der erkannten Kuh waren, auf dem Bild erschienen. Zusätzlich wurde geprüft, ob der im Bild erkannte Kuhrücken vollständig im Tiefenbild zu sehen war und eine

Mindestanzahl von Punkten enthielt. Nur diese 3D-Aufnahmen (Frames) wurden für die weitere Datenauswertung verwendet. Schlussendlich wurden alle nicht dem Kuhrücken zugeordneten Tiefenwerte verworfen.

4.3 Eruierung von Messpunkten

Für die automatische Erfassung des Locomotion Scores wurde sich an der Methode von SPRECHER et al. (1997) orientiert. Das bedeutet, dass die Wirbelsäule als wichtigstes Merkmal automatisch erkannt werden musste. Der letzte Teil der Wirbelsäule (von den Hüfthöckern bis zum Schwanz), wurde nicht mit betrachtet, da die signifikante Krümmung der Wirbelsäule im ersten Drittel über dem Schulterbereich des Tieres stattfindet.

Es mussten eindeutige Messpunkte in den Tiefenbildern eruiert werden, welche eine automatische Identifizierung der Wirbelsäule ermöglichen. Die Wirbelsäule wurde im Tiefenbild als die Polylinie definiert, welche die höchsten Punkte vom Hüftbereich bis zum Nackenbereich enthält.

4.3.1 Berechnung 1

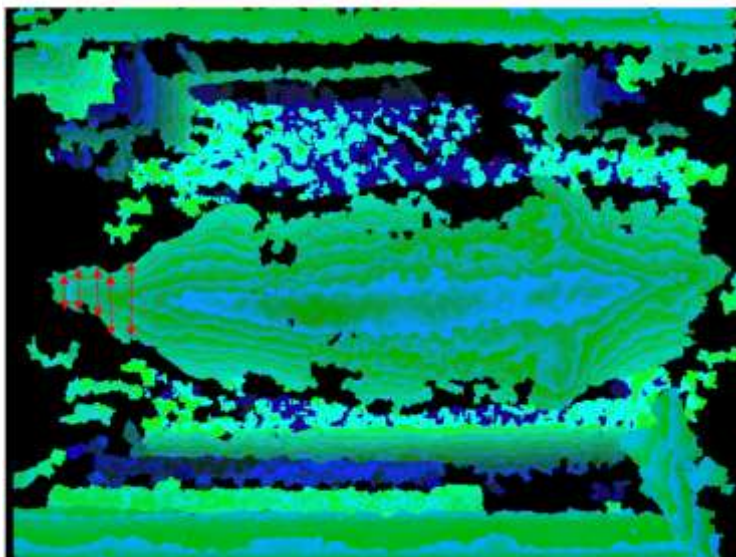
Als erstes wurden die Hüfthöcker identifiziert. Hierzu wurde nur das hintere Drittel des Tiefenbildes betrachtet und die höchsten Punkte am äußeren Rand gesucht. Diese Punkte stellen die Hüfthöcker der Kuh dar. Anschließend wurde zwischen diesen Punkten eine Gerade gezogen (Abbildung 17). In der Mitte der Linie wurde der höchste Punkt gesucht und von diesem ausgehend die Polylinie berechnet. So wurde automatisch die Wirbelsäule identifiziert und nachgezeichnet. Diese Berechnung wurde anhand der Daten beider 3D Sensoren durchgeführt (Abbildung 12) und ein automatischer Score berechnet.



Abbildung 17: 2,5D-Tiefenbild einer Kuh mit algorithmisch identifizierter Hüft- und Wirbelsäulenlinie in Rot (Ansicht von oben). Die Hüftlinie ist die Verbindungsline der beiden höchsten, äußerten Punkte. Die Wirbelsäule wird als Polylinie der höchsten Punkte ausgehend von der Hüftlinie ermittelt.

4.3.2 Berechnung 2

Im zweiten Schritt wurde der Endpunkt der Rückenlinie definiert. Die Position von Hals und Kopf sollte nicht mit bewertet werden, sondern lediglich der Bereich zwischen den Schultern und Hüfthöckern. Hierfür wurde jeweils von der linken Seite des Bildes beginnend die Breite des Tiefenbildes gemessen (Abbildung 18). Im Abstand von zwei Zentimetern wurde die Breite ermittelt. Wiesen vier Messungen hintereinander einen positiven Trend auf, wurde diese Stelle als Endpunkt für die Berechnung der Rückenlinie festgelegt. Für diese Berechnung wurden nur noch die Daten von Sensor 1 herangezogen.



Abstandsmessung

Abbildung 18: Berechnung des Endpunkts der Rückenlinie. Abstandsmessung der Breite des Kuhhalses im Abstand von 2 cm. Weisen vier Messungen hintereinander positive Werte auf, wird diese Stelle als Endpunkt definiert.

4.4 Berechnung der automatischen Scores

Die Berechnung des Locomotion Scores erfolgte mit der 3D-Software Final Surface® (Gesellschaft zur Förderung angewandter Informatik e.V.). Zur Berechnung des automatischen Locomotion Scores wurde die Rückenlinie aus dem 2,5D-Tiefenbild in den 3D-Raum überführt. An die Linie wurde ein Kreis approximiert, dessen Radius umgekehrt proportional zu der Krümmung der Kuhrückenlinie ist (Abbildung 19). Je kleiner der Radius des Kreises ist, desto gekrümmter ist der Rücken der Kuh. Der Median der Krümmungswerte von jedem erkannten Kuhrücken der Bildsequenz wurde bestimmt.

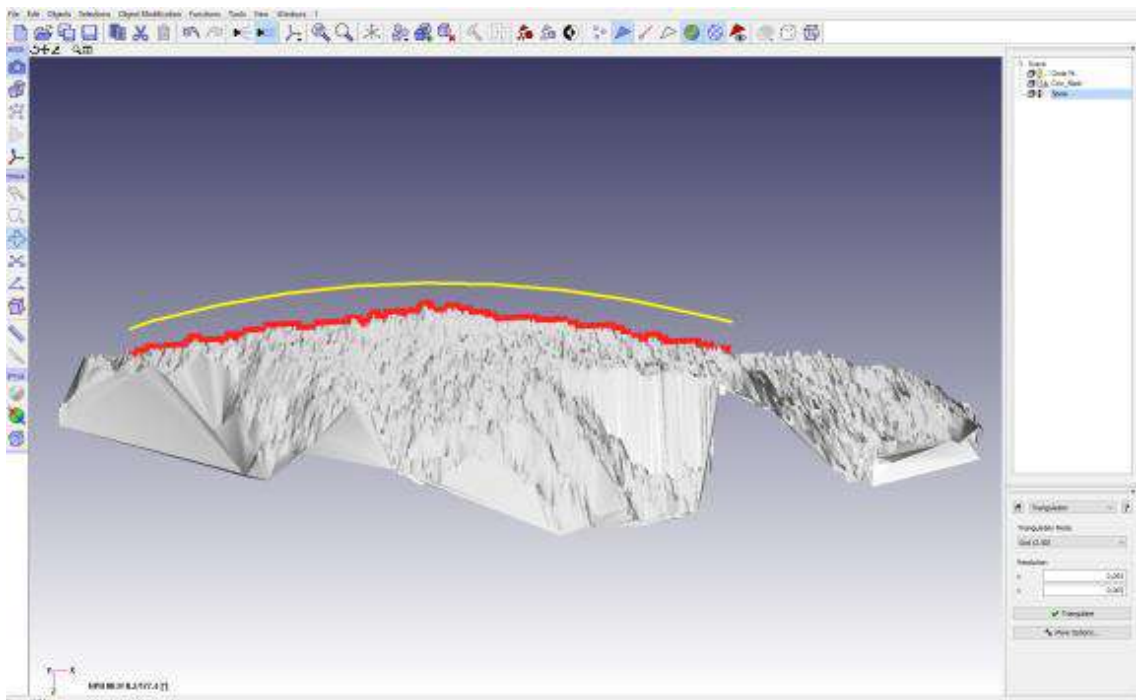


Abbildung 19: 3D-Dreiecknetz einer Kuh mit der gefundenen Rückenlinie in Rot und der approximierten Rückenkrümmung in Gelb (Seitenansicht) in der 3D-Software „Final Surface“

Da eine Bildsequenz aus 1 – 60 Einzelbildern bestand, wurde der Krümmungswert für jedes Einzelbild bestimmt und für die Bildsequenz als Median ausgegeben. Zur Klassifizierung wurden für jeden Score Beispieldühe visuell ausgewählt und anhand der Krümmungsgrade dieser Kühe eine erste Klassifizierung vorgenommen. Die Krümmungsgrade bewegten sich zwischen 0,1 und 0,9. Die Klassen für die fünf Scores wurden anhand der Beispieldühe wie in Tabelle 3 festgelegt.

Tabelle 3: Erste Klassifizierung der Krümmungsgrade

Klassengrenzwert Krümmungsgrad	Locomotion Score
< 0,3	1
0,3 – 0,48	2
> 0,48 – 0,6	3
> 0,6 – 0,7	4
> 0,7	5

4.5 Visuelles Locomotion Scoring

Zur Validierung der automatischen Locomotion Scores wurden die Videoaufnahmen der 2D Kamera 2 gesichtet und anhand dieser Zwei-Sekunden-Aufnahmen, welche verlangsamt auf vier Sekunden abgespielt wurden, der Locomotion Score vergeben. Außerdem wurden die Aufnahmen der Kamera 1 angesehen, um besondere Vorkommnisse im Versuchsablauf zu ermitteln. Als Besonderheit wurden Ereignisse wie die Darmentleerung, die Präsenz eines Mitarbeiters, eine verminderte Bewegung oder Staus im Laufgang notiert.

Die Hauptauswertung erfolgte durch die Verfasserin dieser Arbeit, nachfolgend Beobachter 1 genannt. Ein geringerer Anteil der Daten wurde auch von zwei weiteren Beobachtern (Beobachter-Trio) durchgeführt, sodass aus drei Scores ein gemeinsamer Score gebildet werden konnte.

Beobachter 1 hat zur Validierung der Algorithmen 200 Kühe visuell bewertet. Es wurde von jeder Kuh der erste Datensatz angesehen. Davon ausgehend wurde jeweils der mindestens sieben Tage entfernte Datensatz, sofern vorhanden, ausgewertet und von diesem Datensatz erneut der mindestens sieben Tage entfernte Datensatz. Insgesamt standen von Beobachter 1 somit 1933 Datensätze zur Verfügung.

Zum Ende des Projektzeitraums haben auch zwei weitere Beobachter die Aufnahmen angesehen und bewertet. Aus Gründen der Zeiteffizienz wurde nur jeweils die erste Aufnahme jeder Kuh bewertet. Insgesamt wurden so 42 Kühe von drei Beobachtern unabhängig bewertet. Das sind 20 % aller bewerteten Tiere von Beobachter 1.

Für die Auswertungen standen demnach unterschiedliche Stichproben zur Verfügung. Diese sind in Tabelle 4 aufgelistet.

Tabelle 4: Übersicht über die für die Berechnungen und Auswertungen zugrundeliegenden Datensätze und Stichproben

Stichprobe	Anzahl Tiere	Anzahl Datensätze	Anzahl Beobachter	Bemerkungen
1	111	722	1	Erste Stichprobe
2	197	197	1	Zweite Stichprobe Nur erste Bewertung
3	200	1489	1	Stichprobe max.
4	181	1052	1	Stichprobe max. Ohne Besonderheiten
5	181	181	1	Stichprobe max. Ohne Besonderheiten Nur erste Bewertung
6	40	40	3	Stichprobe Beobachter-Trio Ohne Besonderheiten Nur erste Bewertung

Validierung und Klassifizierung der Algorithmen zur Ableitung des Locomotion Scores

4.6 Automatisches Locomotion Scoring

Die Entwicklung des automatischen Locomotion Scores erfolgte in vier Berechnungsschritten. Im Versuchszeitraum vom 03.08.2016 bis 30.12.2016 wurden 5925 Datensätze aufgezeichnet.

Insgesamt wurden 320 Ordner angelegt. In diesen Kuh-Ordnern befanden sich durchschnittlich 20 weitere Ordner (min. = 1; max. = 115). 200 der 320 Kühe wurden ausgewertet.

4.6.1 Auswertung 1

Im ersten Schritt wurden alle visuellen Datensätze mit den Datensätzen des automatischen Systems verglichen. Die automatische Berechnung erfolgte anhand der in Tabelle 3 vorgegebenen Klassifizierung, wobei die Daten aus beiden 3D Sensoren (siehe Abbildung 12) in die Berechnung eingingen.

Tabelle 5: Darstellung der Ergebnisse aus Berechnung 1 (Polylinie der Wirbelsäule ohne definierten Endpunkt) anhand von Stichprobe 1. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores

	N (vLS vs. aLS)	%	N (vLS vs. aLS) (Score 3, 4, 5 = 3)	%
Übereinstimmungen	359	50	459	64
Abweichungen	363	50	263	36
Gesamt	722	100	722	100

Diese Art der Berechnung des Locomotion Scores zeigt eine Übereinstimmung von 50 % mit den visuell vergebenen Scores (Tabelle 5).

Anschließend wurden die Scores 3, 4 und 5 sowohl bei den visuellen als auch bei den automatischen Ergebnissen als 3 zusammengefasst. Dadurch verbesserte sich die Übereinstimmung der automatischen und visuellen Scores auf 64 % (Tabelle 5).

4.6.2 Auswertung 2

Für die zweite Berechnung wurde eine kleinere Stichprobe gewählt. Es wurde nur noch die jeweils erste Bewertung von jeder Kuh ausgewertet. Dadurch verringerte sich die Stichprobe auf 197 Datensätze. Außerdem wurde ein fester Endpunkt der Rückenlinie definiert (siehe 4.3.2 Berechnung 2). Zur Auswertung dieser und aller folgenden Berechnungen wurden nur noch die Bilddaten von Sensor 1 genutzt.

Tabelle 6: Darstellung der Ergebnisse aus Berechnung 2 (Polylinie der Wirbelsäule mit definiertem Endpunkt) anhand von Stichprobe 2. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores - „Auswertung ohne Hals“

	N (vLS vs. aLS) ohne Hals	%	N (vLS vs. aLS) (Score 3, 4, 5 = 3)	%
Übereinstimmungen	110	56	115	58
Abweichungen	87	44	82	42
Gesamt	197	100	197	100

Der Vergleich der visuellen, dreistufigen Scores mit den automatischen, dreistufigen Scores ergab bei der Berechnung ohne Hals für die verringerte Stichprobe eine Übereinstimmung von 58 % (Tabelle 6). Da dieser Wert unter dem zuvor in Berechnung 1 ermittelten Wert von 64 % liegt (Tabelle 5), wurde auch für die kleinere Stichprobe der zweiten Berechnung der in Berechnung 1 angewandte Algorithmus zur

Auswertung mit Hals angewandt. Das Ergebnis zeigt eine Übereinstimmung von 57 % (Tabelle 7).

Tabelle 7: Darstellung der Ergebnisse aus Berechnung 1 (Polylinie der Wirbelsäule ohne definierten Endpunkt) anhand von Stichprobe 2. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores - „Auswertung mit Hals“

	N (vLS vs. aLS) mit Hals	%	N (vLS vs. aLS) (Score 3, 4, 5 = 3)	%
Übereinstimmungen	106	54	112	57
Abweichungen	91	46	85	43
Gesamt	197	100	197	100

Da sich die Ergebnisse der Berechnungen mit und ohne Hals so sehr ähneln, wurde überprüft, wie sich die Häufigkeiten der einzelnen Scores verteilen (Abbildung 20).

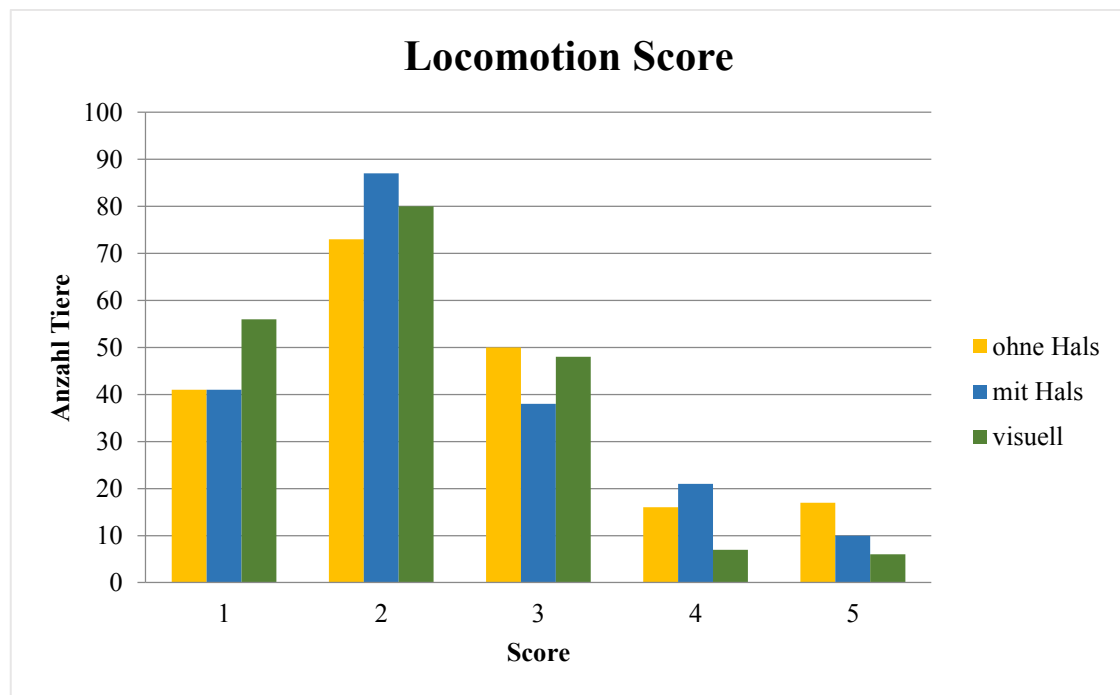


Abbildung 20: Vergleich der Häufigkeiten der einzelnen Scores bei der Berechnung mit Hals und ohne Hals jeweils für die Stichprobe von 197 Datensätzen (Stichprobe 2)

Das Ergebnis zeigt, dass visuell insgesamt mehr Scores mit der Note 1 vergeben wurden als durch das automatische System. Für die Scores 4 und 5 lagen die Häufigkeiten dagegen bei dem automatischen System über denen der visuellen Bonitur.

Um genauer bestimmen zu können, bei welchen visuellen Scores das automatische System zu einem anderen Ergebnis kommt, wurde eine Matrix erstellt, welche die

Anzahl der Abweichungen der automatischen Scores von den visuellen Scores je Größe der Abweichung und je visuellem Score darstellt.

Tabelle 8: Anzahl der Abweichungen der automatischen Scores vom visuellen Score je Größe der Abweichung und je visuellem Score anhand von Stichprobe 2 – „Auswertung mit Hals“ (Berechnung 1)

Ab- weichung (vLS - aLS)	1	2	3	4	5
2	0	0	0	1	0
1	0	8	17	2	1
-1	19	16	13	2	0
-2	4	5	2	0	0
-3	0	1	0	0	0
0	33	50	16	2	5

In der Matrix für die Auswertungen „mit Hals“ (Tabelle 8) ist kein eindeutiger Trend erkennbar. Bei einer visuellen Bewertung mit 1 oder 2 lag das automatische Ergebnis überwiegend eine Note unter dem visuellen Ergebnis, aber besonders wenn visuell eine 3 vergeben wurde, hat das automatische System eine Note höhere vergeben.

Wurden dagegen dieselben Auswertungen bei der automatischen Berechnung „ohne Hals“ durchgeführt (Tabelle 9), zeichnete sich deutlich ab, dass in den meisten Fällen die automatische Berechnung der Scores eine höhere Note ausgab als die visuelle Beurteilung.

Tabelle 9: Anzahl der Abweichungen der automatischen Scores vom visuellen Score je Größe der Abweichung und je visuellem Score anhand von Stichprobe 2 – „Auswertung ohne Hals“ (Berechnung 2)

Ab- weichung (vLS-aLS)	1	2	3	4	5
2	0	0	0	1	0
1	0	8	9	0	0
-1	17	21	10	4	0
-2	6	4	6	0	0
-3	0	1	0	0	0
0	33	46	23	2	6

Daraus konnte geschlossen werden, dass die Klassen der automatischen Berechnung noch nicht optimal festgelegt waren. Um die Grenzwerte zu kontrollieren, wurde die Ausgabe der automatischen Scores erweitert, indem neben den Scores zusätzlich die

tatsächlichen Messwerte gespeichert wurden. Mit dem Statistikprogramm R wurden dann iterativ die optimalen Grenzen zur Unterteilung der Messwerte in Scores berechnet, um eine möglichst hohe Übereinstimmung zwischen den (automatischen) Scores und den visuell validierten Einstufungen zu erreichen. Dies erfolgte, indem jeweils von den zu Beginn festgelegten vier Grenzwerten ein Grenzwert zwischen dem vorherigen und nachfolgenden Grenzwert in Schritten von 0,01 variiert wurde (Bereich Messwerte ca. 0 – 1,2) und für jede dieser Positionen des Grenzwertes die Übereinstimmung zwischen dem visuellen und automatischen Score ermittelt wurde. An der Position, an welcher die höchste Übereinstimmung auftrat, wurde der optimierte Grenzwert definiert. Auf diese Weise wurden alle Grenzwerte für die Unterteilung der Messwerte in automatische Scores optimiert. Die optimierten Klassengrenzen sind in Tabelle 10 dargestellt.

Tabelle 10: Klassifizierung mit optimierten Grenzwerten anhand der Ergebnisse von Berechnung 2

LS	Klassengrenzwerte Krümmungsgrad	Neue Klassengrenzwerte Krümmungsgrad
1	< 0,3	< 0,36
2	0,3 – 0,48	0,36 – < 0,65
3	> 0,48 – 0,6	0,65 – < 0,81
4	> 0,6 – 0,7	0,81 – < 0,94
5	> 0,7	>= 0,94

4.6.3 Auswertung 3

4.6.3.1 Stichprobe 3

Mit den neuen Klassengrenzen wurden die Daten erneut ausgewertet. Als erstes wurde die Übereinstimmung bei der größten Stichprobe (Stichprobe max. = Stichprobe 3) getestet. Diese Stichprobe bestand aus 200 Kühen mit 1489 Messungen, welche durch Beobachter 1 visuell eingestuft wurden. Hier lag die Übereinstimmung bei 54 % bei dem Vergleich der fünf Stufen (Tabelle 11) und bei 63 % für eine 3-stufige Einteilung (Score 3, 4, 5 = 3).

Tabelle 11: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 3. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores

	N (vLS vs. aLS)	%	N (vLS vs. aLS) (Score 3, 4, 5 = 3)	%
Übereinstimmungen	803	54	940	63
Abweichungen	686	46	549	37
Gesamt	1489	100	1489	100

Zur besseren Veranschaulichung werden die Messwerte der Rückenkrümmung zu den jeweiligen visuellen Scores in Abbildung 21 dargestellt. Hier fallen besonders die hohen Werte über $> 1,2$ auf. Um mögliche Messfehler in der Auswertung zu vermeiden, wurden alle Datensätze gefiltert, die in der Spalte „Besonderheiten“ eine Bemerkung hatten sowie Datensätze, deren Messwerte größer als 1,2 sind. Außerdem wurden Datensätze entfernt, bei denen die Abweichung des visuellen und des automatischen Scores > 1 ist (insgesamt 56).

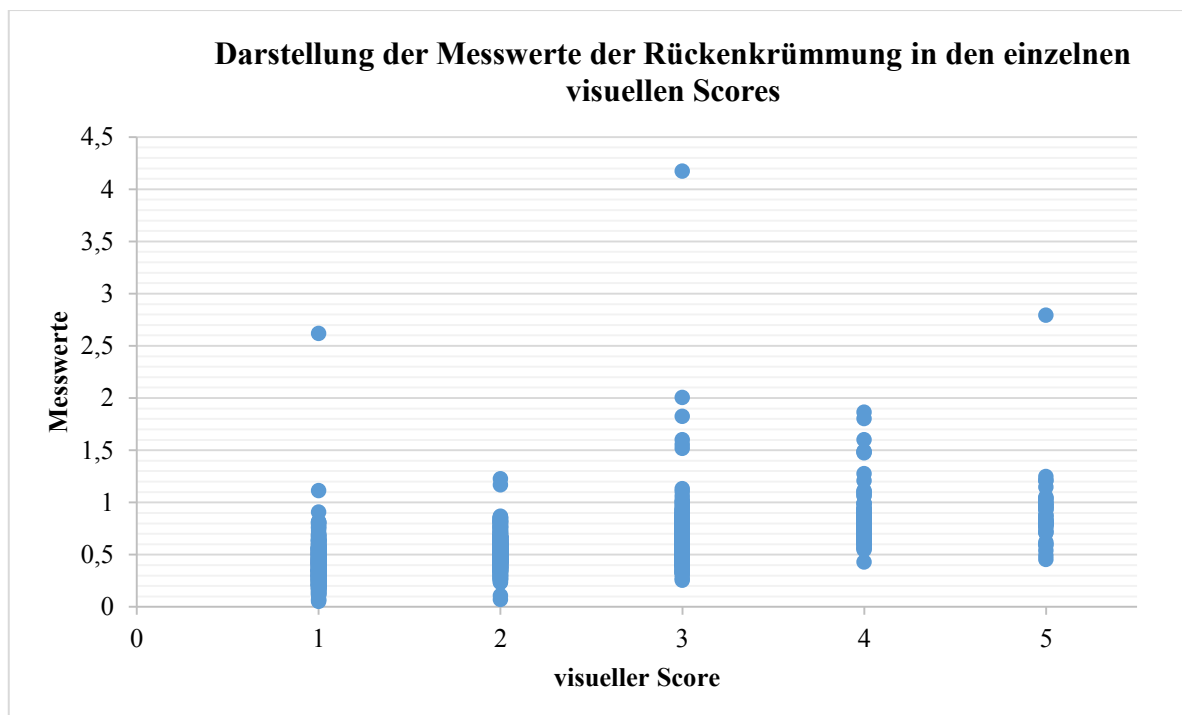


Abbildung 21: Darstellung der Messwerte der Rückenkrümmung aus Berechnung 3 zu den einzelnen visuellen Scores aus Stichprobe 3

4.6.3.2 Stichprobe 4

Mit dieser Stichprobe wurde die Übereinstimmung der visuellen und automatischen Scores überprüft, welche für das Modell mit fünf Stufen 59 % ergab und für das mit drei Stufen 65 % (Tabelle 12).

Tabelle 12: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 4. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores

	N (vLS vs. aLS)	%	N (vLS vs. aLS) (Score 3, 4, 5 = 3)	%
Übereinstimmungen	616	59	688	65
Abweichungen	436	41	364	35
Gesamt	1052	100	1052	100

Die Verteilung der Messwerte in den dazugehörigen visuellen Scores ist in Abbildung 22 in einem Box-Plot dargestellt. Hier ist sehr gut zu sehen, dass die Messwerte der Rückenkrümmung zwar relativ weit streuten, 50 % der Werte jedoch in gut abgrenzbaren Gruppen lagen (Box).

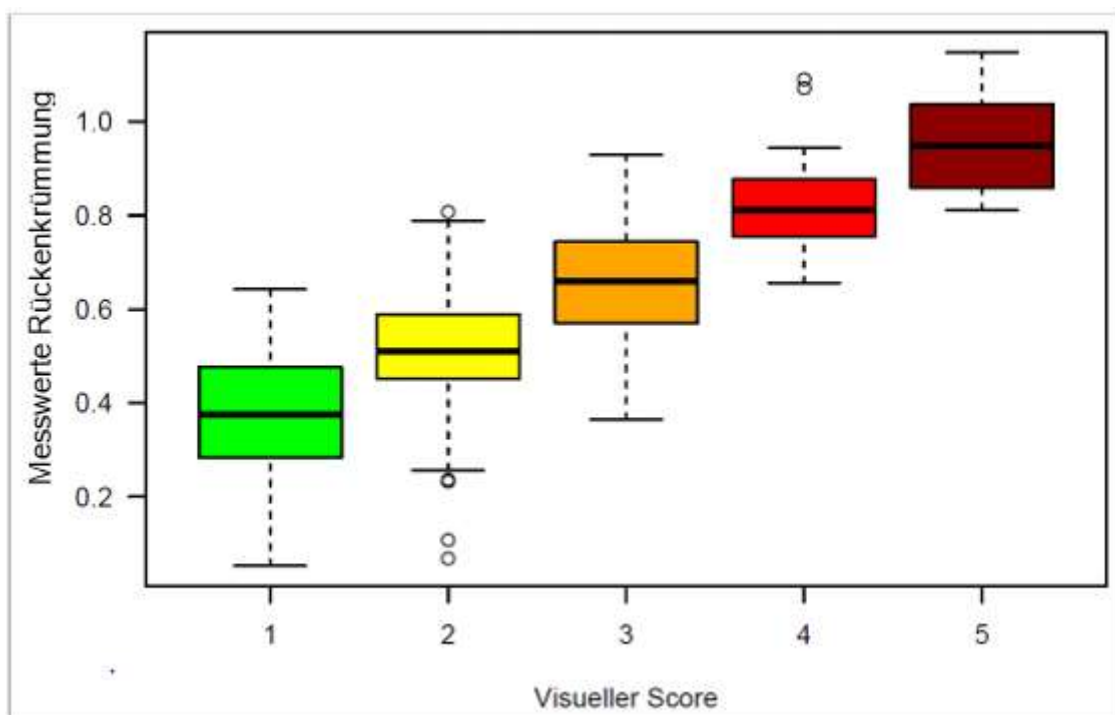


Abbildung 22: Box-Plot der Messwerte der Rückenkrümmung aus Berechnung 3 zu den visuellen Scores anhand von Stichprobe 4

Eine Berechnung der Korrelation nach Spearman hat $r_s = 0,73$ ergeben (Tabelle 13). Die Varianzanalyse ANOVA hat außerdem gezeigt, dass der visuelle Score den Messwert

der Rückenkrümmung signifikant beeinflusst. Im Tukey HSD Test wurde ein signifikanter Unterschied der Messwerte in den einzelnen Scores bestätigt ($p < 0,05$).

Tabelle 13: Statistische Auswertung von Berechnung 3 anhand von Stichprobe 4

Korrelation					0,7269103
(visueller Score, automatischer Score, Methode = „Spearman“)					
Varianzanalyse ANOVA					
(Zielvariable: Messwerte Rückenkrümmung, Faktor: visueller Score)					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
Faktor (Visueller Score)	4	19,672	4,9180	357,37	<2,2e-16 ***
Residuals	1047	14,409	0,0138		
Signif. Codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1					
Tukey HSD					
Tukey multiple comparison of means					
95% family-wise confidence level					
Fit: aov(Messwerte Rückenkrümmung – visueller Score)					
	diff	lrw	upr	P adj	
2-1	0,1343308	0,10958350	0,1590781	0,0000000	
3-1	0,2736735	0,24735342	0,2999935	0,0000000	
4-1	0,4337062	0,39219347	0,4752189	0,0000000	
5-1	0,5668078	0,47920400	0,6544117	0,0000000	
3-2	0,1393427	0,11412849	0,1645568	0,0000000	
4-2	0,2993754	0,25855485	0,3401959	0,0000000	
5-2	0,4324770	0,34519906	0,5197550	0,0000000	
4-3	0,1600327	0,11823999	0,2018254	0,0000000	
5-3	0,2931344	0,20539749	0,3808712	0,0000000	
5-4	0,1331016	0,03967617	0,2265271	0,0009964	

Außerdem wurde ein Test auf Selektivität und Sensitivität durchgeführt. Hierfür wurden alle Tiere mit einem Score von 1 als „gesund“ definiert und alle Tiere mit einer schlechteren Note als „krank“. Die visuelle Bewertung galt als wirklicher Zustand, der automatische Score als Testergebnis (Formeln siehe Anhang). Bei Stichprobe 4 hatte der automatische Score eine Sensitivität von 96 % und eine Spezifität von 48 %. Die Genauigkeit lag bei 82 % (Tabelle 14).

Tabelle 14: Test auf Selektivität und Sensitivität von Berechnung 3 anhand von Stichprobe 4

Kriterium	Definition	Anzahl der Kühe (Σ 1052)
Richtig positiv	Lahme Kuh als lahm erkannt	714
Falsch positiv	Gesunde Kuh als lahm erkannt	159
Richtig negativ	Gesunde Kuh als gesund erkannt	148
Falsch negativ	Lahme Kuh als gesund erkannt	31
Auswertung		
Sensitivität (krank = krank)		96 %
Spezifität (gesund = gesund)		48 %
Positiver prädikativer Wert (Fehlerrate)		18 %
Negativer prädikativer Wert (Genauigkeit)		82 %
Falsch positiver Wert		3 %
Falsch negativer Wert		15 %

4.6.3.3 Stichprobe 5

In der nächsten Analyse wurde nur noch der erste vergebene Score pro Kuh betrachtet, um sowohl Beeinflussungen von Beobachter 1 durch die zuvor gegebene Bewertung zu minimieren als auch den Einfluss der unterschiedlichen Häufigkeiten, mit denen die Kühe in die Bewertung eingehen, zu schmälern. In dieser Auswertung wurden ebenfalls alle Tiere, welche eine Bemerkung bei „Besonderheiten“ hatten oder deren Differenz zwischen automatischem und visuellen Score > 1 war, ausgeschlossen. Als Stichprobe 5 blieben somit 181 Kühe mit 181 Datensätzen. Es wurde erneut die genaue Übereinstimmung für fünf und drei Stufen geprüft. Diese lag bei 56 % bzw. 64 % (Tabelle 15).

Tabelle 15: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 5. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores

	N (vLS vs. aLS)	%	N (vLS vs. aLS) (Score 3, 4, 5 = 3)	%
Übereinstimmungen	102	56	116	64
Abweichungen	79	44	65	36
Gesamt	181	100	128	100

Verglichen mit Stichprobe 4 ist keine große Veränderung eingetreten. Zur besseren Veranschaulichung der Messergebnisse wurde auch hier ein Box-Plot-Diagramm erstellt. In diesem zeigt sich, dass die Messergebnisse in den Scores 4 und 5 sehr dicht beieinander lagen (Abbildung 23). Die Messwerte von Score 5 wurden vollständig von Score 4 abgedeckt.

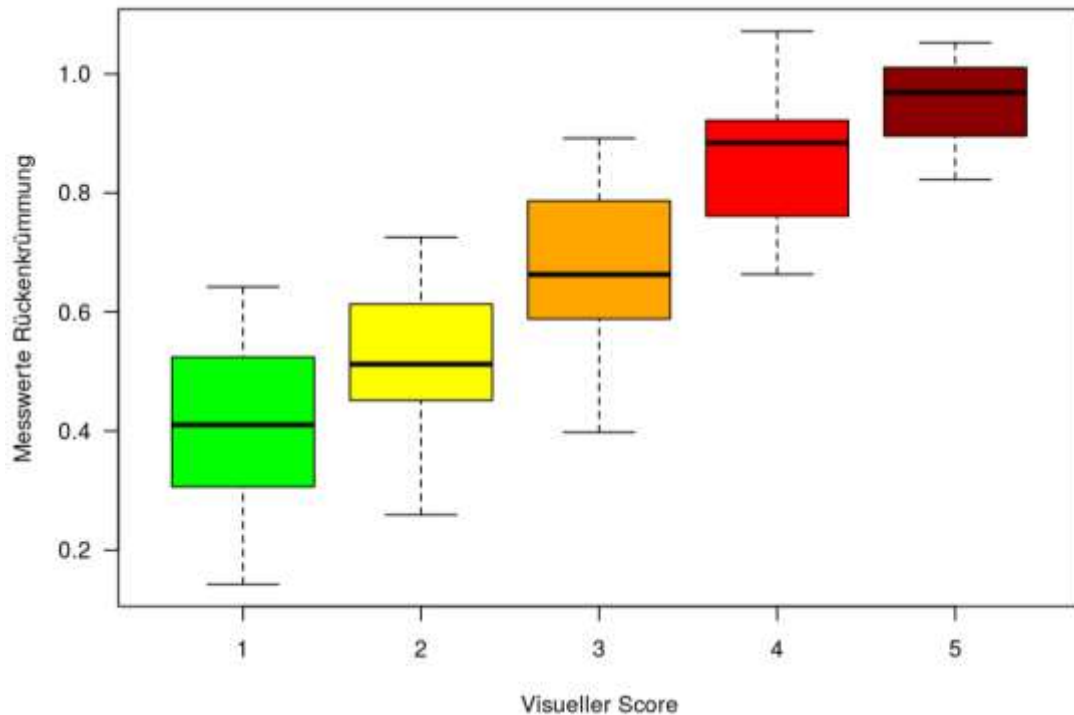


Abbildung 23: Box-Plot der Messwerte der Rückenkrümmung aus Berechnung 3 zu den visuellen Scores anhand von Stichprobe 5

In der statistischen Auswertung ergab die Korrelation nach Spearman den Wert $r_s = 0,70$ und war somit etwas geringer als bei Stichprobe 4, aber der Unterschied zwischen den Messwerten der Rückenkrümmung und den Scores der visuellen Beurteilung war noch immer hoch signifikant ($p < 0,001$). Der Tukey HSD Test bestätigte die Vermutung, dass die Stufen 4 und 5 nicht signifikant unterschiedlich sind ($p = 0,757$) (Tabelle 16).

Tabelle 16: Statistische Auswertung von Berechnung 3 anhand von Stichprobe 5

Korrelation						0,7026146
(visueller Score, automatischer Score, Methode = „Spearman“)						
Varianzanalyse ANOVA						
(Zielvariable: Messwerte Rückenkrümmung, Faktor: visueller Score)						
	DF	Sum Sq	Mean Sq	F value	Pr (>F)	
Faktor (Visueller Score)	4	3,1182	0,77955	52,294	< 2,2e-16 ***	
Residuals	176	2,6236	0,01491			
Signif. Codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1						
Tukey HSD						
Tukey multiple comparison of means						
95% family-wise confidence level						
Fit: aov(Messwerte Rückenkrümmung – visueller Score)						
	diff	lwr	upr	P adj		
2-1	0,11116028	0,05001724	0,1723033	0,0000129		
3-1	0,26499093	0,19746974	0,3325121	0,0000000		
4-1	0,44016685	0,31900440	0,5613293	0,0000000		
5-1	0,53703130	0,33741507	0,7366475	0,0000000		
3-2	0,15383065	0,08977442	0,2178869	0,0000000		
4-2	0,32900657	0,20974036	0,4482728	0,0000000		
5-2	0,42587101	0,22740004	0,6243420	0,0000002		
4-3	0,17517592	0,05251757	0,2978343	0,0011142		
5-3	0,27204036	0,07151264	0,4725681	0,0022947		
5-4	0,09686444	-0,12748483	0,3212137	0,7571082		

Zudem wurde überprüft, wie die Übereinstimmung zwischen automatischen und visuellen Scores innerhalb der einzelnen visuellen Scores ist (Tabelle 17). Es zeigte sich, dass die Übereinstimmung zwischen vLS und aLS in Locomotion Score 2 am höchsten ist, nämlich 80 %. Die geringste Übereinstimmung war in Locomotion Score 1 mit 39 % vorhanden.

Auch mit Stichprobe 5 wurde ein Test auf Genauigkeit durchgeführt. 95 % wurden richtig als lahm erkannt und 39 % wurden korrekt als gesund eingestuft. Die Genauigkeit liegt bei 79 % (Tabelle 18).

Ergebnisse

Tabelle 17: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 5. Anzahl der Übereinstimmungen und Abweichungen der visuellen und automatischen Scores innerhalb der visuellen Scores

		Übereinstimmungen	Abweichungen	Gesamt
vLS 1	N (vLS vs. aLS)	21	33	54
	%	39	61	100
vLS 2	N (vLS vs. aLS)	55	14	69
	%	80	20	100
vLS 3	N (vLS vs. aLS)	20	26	46
	%	43	57	100
vLS 4	N (vLS vs. aLS)	4	5	9
	%	44	56	100
vLS 5	N (vLS vs. aLS)	2	1	3
	%	67	33	100

Tabelle 18: Test auf Spezifität und Sensitivität von Berechnung 3 anhand von Stichprobe 5

Kriterium	Definition	Anzahl der Kühe (Σ 181)
Richtig positiv	Lahme Kuh als lahm erkannt	121
Falsch positiv	Gesunde Kuh als lahm erkannt	33
Richtig negativ	Gesunde Kuh als gesund erkannt	21
Falsch negativ	Lahme Kuh als gesund erkannt	6
Auswertung		
Sensitivität (krank = krank)		95 %
Spezifität (gesund = gesund)		39 %
Positiver prädikativer Wert (Fehlerrate)		21 %
Negativer prädikativer Wert (Genauigkeit)		79 %
Falsch positiver Wert		3 %
Falsch negativer Wert		18 %

4.6.3.4 Stichprobe 6

Um die Subjektivität der visuellen Bewertung zu verringern, wurde von zwei weiteren Personen die Rückenkrümmung der Kühe bewertet. Es wurde von vornherein nur eine Aufnahme pro Kuh angesehen. Auch hier wurden am Ende wieder die Kühe mit „Besonderheiten“ und die, bei denen der automatische und der visuelle Score um mehr als eine Note differieren, aus der Stichprobe entfernt. Für Stichprobe 6 haben dementsprechend drei Beobachter 40 Kühe mit 40 Datensätzen ausgewertet. Da bei drei Bewertungen die Score-Noten nicht mehr nur ganze Noten sind, wurden die visuellen Scores der Beobachter gerundet und anschließend auf Übereinstimmung mit den automatischen Scores überprüft. Eine Übereinstimmung von 67 % bzw. 72 % wurde im fünfstufigen bzw. dreistufigen Modell erreicht (Tabelle 19).

Tabelle 19: Darstellung der Ergebnisse aus Berechnung 3 anhand von Stichprobe 6. Anzahl der Übereinstimmungen und Abweichungen der visuellen und der automatischen Scores

	N (vLS vs. aLS)	%	N (vLS vs. aLS) (Score 3, 4, 5 = 3)	%
Übereinstimmungen	27	67	29	72
Abweichungen	13	33	11	28
Gesamt	40	100	40	100

Aufgrund der kleineren Stichprobe reichten die visuellen Scores nur von 1 – 3. Die automatischen Scores reichten dagegen bis 4. Die Verteilung der Messwerte zu den einzelnen Scores (hier ungerundet) ist Abbildung 24 zu entnehmen. Die größte Streuung an Messwerten war bei einem visuellen Score von 1,67 vorhanden. Diese Messwerte streuten nahezu über alle Messwerte, die den visuellen Scores von 1,0 – 2,67 entsprechen.

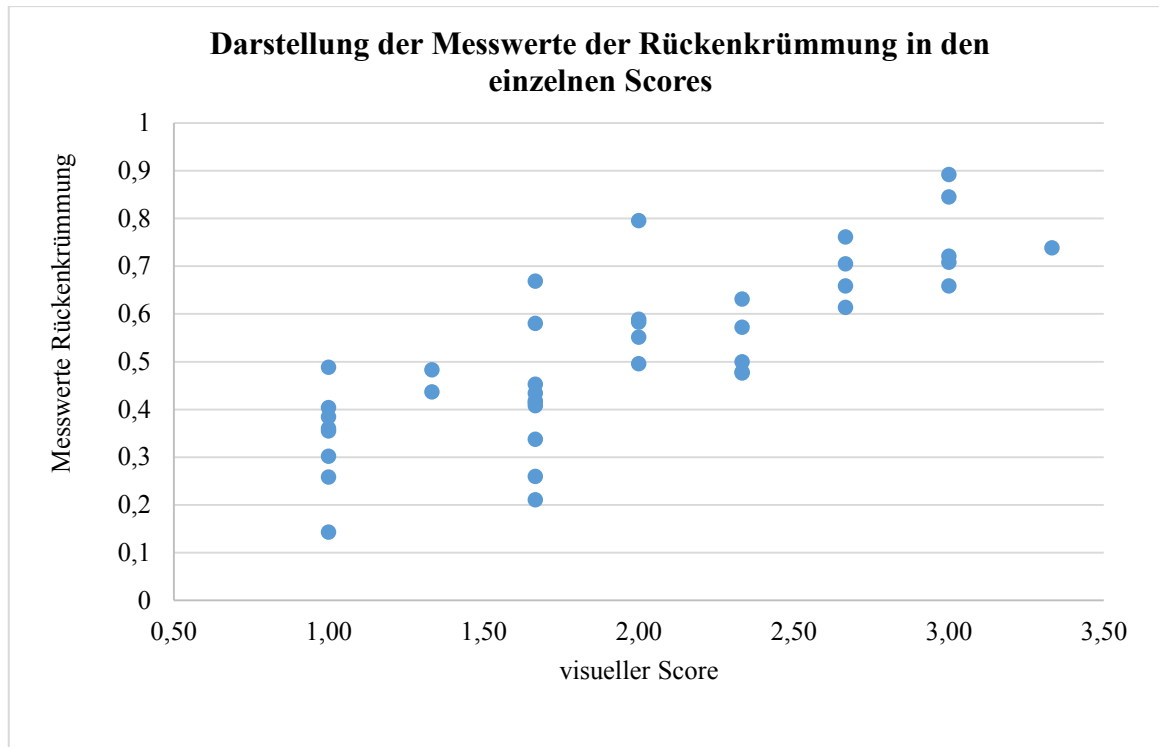


Abbildung 24: Darstellung der Messwerte der Rückenkrümmung aus Berechnung 3 in den einzelnen Scores anhand von Stichprobe 6

Die Korrelation bei dieser Stichprobe war sehr stark ($r_s = 0,73$) und die Unterschiede der Messwerte in den einzelnen Scores waren hoch signifikant ($p < 0,001$). Auch der Tukey HSD Test bestätigte die Unterschiede der Messwerte in den Scores (Tabelle 20).

Tabelle 20: Statistische Auswertung von Berechnung 3 anhand von Stichprobe 6

Korrelation					0,7317972
(visueller Score, automatischer Score, Methode = „Spearman“)					
Varianzanalyse ANOVA					
(Zielvariable: Messwerte Rückenkrümmung, Faktor: visueller Score)					
	DF	Sum Sq	Mean Sq	F value	Pr (>F)
Faktor (Visueller Score)	2	0,70806	0,35403	24,923	1,396e-07 ***
Residuals	37	0,52558	0,01420		
Signif. Codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1					
Tukey HSD					
Tukey multiple comparison of means					
95% family-wise confidence level					
Fit: aov(Messwerte Rückenkrümmung – visueller Score)					
	diff	lwr	upr	P adj	
2-1	0,131264	0,01856599	0,2439620	0,0193153	
3-1	0,368742	0,23860954	0,4988745	0,0000001	
3-2	0,237478	0,12477999	0,3501760	0,0000264	

Ergebnisse

Der Test für die korrekte Klassifizierung ergab eine Sensitivität von 100 % und eine Spezifität von 40 % (Tabelle 21). Die Fehlerrate lag bei 37 %.

Tabelle 21: Test auf Spezifität und Sensitivität von Berechnung 3 anhand von Stichprobe 6

Kriterium	Definition	Anzahl der Kühe (Σ 40)
Richtig positiv	Lahme Kuh als lahm erkannt	20
Falsch positiv	Gesunde Kuh als lahm erkannt	12
Richtig negativ	Gesunde Kuh als gesund erkannt	8
Falsch negativ	Lahme Kuh als gesund erkannt	0

Auswertung		
Sensitivität (krank = krank)		100 %
Spezifität (gesund = gesund)		40 %
Positiver prädikativer Wert (Fehlerrate)		37 %
Negativer prädikativer Wert (Genauigkeit)		63 %
Falsch positiver Wert		0 %
Falsch negativer Wert		30 %

5. Diskussion

Prozessentwicklung

5.1 Experimentelles Design

Der Durchführungsort für das Locomotion Scoring wurde entsprechend der im Testbetrieb vorhandenen Gegebenheiten ausgewählt. Als ausreichend langer, gerader Laufweg der Kühe, mit der Möglichkeit zur gleichzeitigen Beobachtung von der Seite, kamen nur zwei Wege in Frage. Zum einen der Treibgang zum Vorwarte Hof und zum anderen der Rücktriebweg hinter dem Melkstand. Eine Installation des Systems im Treibgang zum Vorwarte Hof entspräche einem Durchlaufen des automatischen Systems wie bei RAJKONDAWAR et al. (2006) vor dem Melken. Da der Rücktriebweg jedoch bessere Möglichkeiten zum Aufbau der Technik und zur Durchführung der visuellen Bonitur der Kühe bot, wurde sich entsprechend der Beschreibungen in der übrigen Literatur (MANSON UND LEAVER 1988, SPRECHER et al. 1997, WHAY et al. 1997, WINCKLER UND WILLEN 2001) beim experimentellen Design für diesen Treibgang hinter dem Melkstand entschieden.

In der ersten Versuchsphase wurde das visuelle Locomotion Scoring während des Melkens am Nachmittag vorgenommen. Wie in Abbildung 11 dargestellt, positionierte sich die Beobachterin im Mitarbeitergang hinter der letzten Wandöffnung. Bei diesem Versuch stellte sich jedoch schnell heraus, dass die Kühe nicht wie geplant zügig und gleichmäßig an der Öffnung vorbeigehen und von der Beobachterin visuell bewertet werden können. Stattdessen blieben die Kühe stehen, schauten die Beobachterin neugierig an und verzögerten den gesamten Produktionsablauf. Aus diesem Grund wurden in der ersten Versuchsphase weitere Positionen für die visuelle Bewertung getestet. Dabei stellte sich heraus, dass auch eine Positionierung hinter den Kühen (siehe Abbildung 11 Position 2 der Beobachterin) von diesen wahrgenommen wird und zu einer immensen Störung führt. Schlussendlich bot die dritte Position der Beobachterin den besten Blickwinkel für die visuelle Bonitur der Kühe. Die Öffnung im Laufgang war ausreichend groß, sodass eine Schrittfolge beobachtet werden konnte. Allerdings haben die Kühe auch hier die Beobachterin wahrgenommen und den Betriebsablauf gestört. Nach dieser ersten Versuchsphase wurde deutlich, dass eine visuelle Beobachtung durch eine Person, wie in der Literatur beschrieben (MANSON UND LEAVER 1988, SPRECHER et al. 1997, WHAY et al. 1997, WINCKLER UND WILLEN

2001, VAN HERTEM et al. 2017, HANSEN et al. 2018), unter Praxisbedingungen nicht möglich ist, da die Tiere sich nicht natürlich bewegen und der Produktionsprozess erheblich gestört wird. Diese Vorgehensweise stellte sich als nicht praxistauglich heraus. Als Schlussfolgerung wurde das Versuchsdesign geändert. Zur visuellen Beobachtung wurden wie bei FLOWER et al. (2005), BAHR et al. (2008), SONG et al. (2008), POURSAHERI et al. (2010), PLUK et al. (2012) und VIAZZI et al. (2014) Videokameras angeschafft und die Position des 3D Sensors wurde vorverlegt, vom Ende der Versuchsumgebung auf Höhe der Mitte der großen Öffnung (siehe Abbildung 12).

Die Verlegung des 3D Sensors dichter an den Ausgang des Melkstands hatte zur Folge, dass die Kühe erst eine kurze Strecke zurückgelegt haben bis sie an den Kameras entlangliefen. Die Befürchtung war, dass die Bewegung der Kuh noch nicht gleichmäßig sein könnte. Diese Annahme bestätigte sich während der Versuche jedoch nicht. Die Kühe zeigten eine gleichmäßige Bewegung, sofern sie nicht von anderen Kühen behindert wurden.

Kamera 1 wurde so positioniert, dass sie den Laufgang im großen Durchbruch aufnahm (siehe Abbildung 12). Dadurch wurden sowohl vom 3D Sensor 1 als auch von Kamera 1 derselbe Bereich und Moment aufgezeichnet. Um dem Nachteil, dass eine Bewertung der Kühe anhand von Videoaufnahmen nur eine Momentaufnahme ist (BENCH et al. 1974), entgegenzuwirken, wurde eine zweite Kamera am Anfang des Treibgangs montiert. So konnte das Geschehen im gesamten Gang verfolgt und Besonderheiten vermerkt werden. Bei der Auswertung der Videodaten stellte sich diese Kamera als großer Vorteil heraus.

Bei anderen in der Literatur beschriebenen Versuchen wurden die Videokameras zur Aufzeichnung der Daten für visuelle Locomotion Scores in 3,5 m bis 8,0 m Entfernung aufgestellt (FLOWER et al. 2005, BAHR et al. 2008, SONG et al. 2008, POURSAHERI et al. 2010, PLUK et al. 2012, VIAZZI et al. 2014). Aufgrund der baulichen Gegebenheiten vor Ort, konnte die Kamera in diesem Versuch nur 2,5 m vom Versuchsgang entfernt befestigt werden. Dadurch war das Sichtfeld für die visuelle Vergabe der Locomotion Scores kleiner als bei anderen Versuchen.

Die Literatur zeigt auch, dass umso dichter die Kamera am Versuchsgang steht, diese entsprechend höher angebracht ist. Bei einer Entfernung der Kamera von 6 - 8 m zum Versuchsgang reicht eine Höhe von 1,5 m (BAHR et al. 2008, POURSAHERI et al. 2010),

bei einer kürzeren Distanz von 3,5 m wird die Kamera dagegen in 2,5 m Höhe montiert (PLUK et al. 2012, POURSAHERI et al. 2010). Dementsprechend wurde auch im Versuchsdesign des Dissertationsprojekts die Kamera möglichst dicht unter der Decke befestigt, sodass Kamera 1 aus einer Höhe von 2,2 m Videos aufzeichnete. Die nicht zu hohe Montage der Kamera hatte einen positiven Effekt, denn sie führte zu einem besseren Blickwinkel auf die Kuh. Selbstverständlich kann die Rückenkrümmung am besten beurteilt werden, wenn diese direkt auf Rückenhöhe der Kuh betrachtet wird. Da dies technisch jedoch in den meisten Betrieben nicht umsetzbar ist, sollte die Kamera so hoch wie nötig aber so niedrig wie möglich befestigt werden, damit aussagekräftige Beurteilungen vorgenommen werden können.

Die verwendete Technik ist mit derjenigen aus anderen Untersuchungen vergleichbar. Wie bei JABBAR et al. (2017) wurde der ASUS xTion Pro Live als 3D Sensor verwendet. VIAZZI et al. (2014), VAN HERTEM et al. (2014, 2017) und HANSEN et al. (2018) haben die 3D Kinect Kamera verwendet. Ein Vorteil der ASUS xTion Pro Live ist, dass die Kamera kleiner ist, obwohl die gleichen Sensoren verwendet werden wie in der Kinect. Außerdem wurden zum Zeitpunkt der Versuche dieses Projekts die Kinect Kameras bereits nicht mehr hergestellt, sodass eine Vermarktung mit dieser Kamera nicht möglich wäre. Beide 3D Sensoren haben eine Auflösung von 640 x 480 Pixel. VIAZZI et al. (2014) und VAN HERTEM et al. (2014, 2017) haben mit dem 3D Sensor aus einer Höhe von 3,15 m - 3,45 m gemessen, JABBAR et al. (2017) sogar in 3,69 m Höhe. Im vorliegenden Versuch wurde das System dagegen nur in 2,8 m Höhe montiert. Eine höhere Montage wäre in diesem Fall zwar möglich gewesen, jedoch nicht notwendig, da nur eine Schrittfolge angesehen wurde anstatt zwei. Dadurch ist der Versuchsaufbau mit dem von HANSEN et al. (2018) vergleichbar, die den 3D Sensor in nur 2,3 m befestigt haben.

Zu Versuchsbeginn war zur Berechnung des automatischen Scores eine möglichst lange Strecke vorgesehen, die mindestens zwei Schrittfolgen abbildet. Dafür wurden zwei 3D Sensoren hintereinandergeschaltet. Jedoch zeigten die Videoaufnahmen von Kamera 2, dass es im Versuchsgang produktionsbedingt täglich zu Staus kommt. Diese entstehen, wenn eine andere Gruppe zum Melken in den Vorwartehof getrieben wird. Dann wird am Ende des Versuchsgangs ein Tor geschlossen und die Kühe, welche den Melkstand verlassen, stauen sich im Versuchsgang. Dadurch konnte es vorkommen, dass die Tiererkennung eine Kuh-RFID erkannte und die Videodaten dieser Kuh von Sensor 1

korrekt abgespeicherte, die Kuh aber nicht weiterlaufen konnte als bis zur Tiererkennung und somit die Daten von Sensor 2 eine andere Kuh zeigen, die sich nicht bewegt, sondern stillsteht (im Stau). Diese Problematik beschreiben auch VAN HERTEM et al. (2017). In ihren Versuchen unter Praxisbedingungen konnten 38 % der Aufnahmen nicht ausgewertet werden. Davon wurden 81 % aufgrund von Staus im Versuchsgang aussortiert. Da diese Problematik auch in diesem Projekt häufiger beobachtet wurde, wurde entschieden, dass für die erste Entwicklung der Algorithmen und deren Validierung in diesem Versuchsbetrieb nur die Daten von Sensor 1 herangezogen werden. Eine Ausweitung der Sensordaten durch Triggern von zwei Sensoren ist für weiterführende Untersuchungen in anderen Betrieben geplant.

Sowohl die 2D Kameras als auch die 3D Sensoren nahmen 30 Bilder pro Sekunde auf. Dies entspricht den meisten verwendeten Systemen in vergleichbaren Untersuchungen (BAHR et al. 2008, POURSAHERI et al. 2010, VIAZZI et al. 2014, VAN HERTEM et al. 2014 und 2017, HANSEN et al. 2018). Es wurden aber auch schon Systeme mit 60 (FLOWER et al. 2005), 20 (PLUK et al. 2012) oder 15 Bildern pro Sekunde (SONG et al. 2008) verwendet. In jenen Untersuchungen wurde jedoch ausschließlich mit 2D Kameras zur Bonitur gearbeitet.

Die Auflösung der hier verwendeten 2D Kameras ist, gegenüber der anderer Versuche, höher. Mit 1280 x 1024 Grauwert-Pixeln liefern die Kameras deutlich bessere Aufnahmen als die üblicherweise verwendeten Kameras mit einer Auflösung von 1024 x 768 Pixeln (SONG et al. 2008, POURSAHERI et al. 2010, PLUK et al. 2012). Es wird davon ausgegangen, dass durch die höhere Auflösung ein schärferes Bild entsteht und dementsprechend die visuelle Bonitur anhand der Videodaten genauer ausfällt. Außerdem wurden die Bilder als schwarz-weiß-Aufnahmen aufgezeichnet, um weniger Speicherkapazität zu verbrauchen.

Die 3D Sensoren wurden in staubdichten Aluminiumgehäusen über der Stalldecke befestigt. Da der 3D Sensor senkrecht nach unten ausgerichtet war, sammelte sich vor der Linse bzw. vor der Schutzscheibe kein Staub. Aufgrund der Positionierung oberhalb der Decke kam es auch nicht zu Verschmutzungen aus dem Reinigungsprozess des Treibgangs. Auch HANSEN et al. (2018) haben ihre Technik entsprechend in einem IP66 geschütztem Gehäuse angeordnet. Höhe und Art der Montage der Technik erwiesen sich über die Versuchszeit von sechs Monaten als praxistauglich.

Wie bei PLUK et al. (2012), VAN HERTEM et al. (2014, 2017) und HANSEN et al. (2018) wurde mit einer Triggerfunktion gearbeitet. Wenn eine Kuh mit ihrer RFID durch die Antennen lief, wurde die RFID erkannt und die Abspeicherung der Daten von den 2D und 3D Geräten aktiviert. Der Versuchsaufbau sah vor, dass zwei Antennen rechts und links im Gang installiert werden, um zu garantieren, dass jede Kuh erkannt wird. Aufgrund von Störwellen des Radio-Funks konnten jedoch nicht beide Antennen angeschlossen werden, da sich der Messabstand dann erheblich verringert hätte. Es wurden aber beide Platten angebracht, damit sich der Treibgang von 1,28 m auf 0,95 m verschmälert und eine automatische Tiervereinzelung stattfindet. Aufgrund des beobachteten Tierverhaltens wurde die rechte Antenne zur Tiererkennung aktiviert. Diese erkannte die RFID mit einem maximalen Abstand von 40 cm und speicherte sie mit dem Datum und der Uhrzeit der Erkennung ab. Als optimale Position der Tiererkennung wurde der Bereich zwischen den beiden 3D Sensoren ermittelt. So wurden bei Erkennung eines Tieres von Kamera 1 und 2 sowie von Sensor 1 jeweils die letzten 60 Bilder aus dem Ringbuffer gespeichert und von Sensor 2 die nächsten 60 Bilder. Dadurch war sichergestellt, dass die Bildgrundlage von Kamera 1, welche zur visuellen Bonitur genutzt wurde, dem selben Moment entsprach wie die Bildgrundlage von Sensor 1 zur automatischen Berechnung des Locomotion Scores. Anhand desselben Zeitstempels waren die Daten einander zuzuordnen. Mit einer solchen Technik haben JABBAR et al. (2017) noch nicht gearbeitet. Sie haben Zeit und Nummer der Tiere manuell dokumentiert, um diese anschließend den Zeitstempeln der Bildsequenzen zuordnen zu können.

Insgesamt wurden 320 Ordner innerhalb der Projektlaufzeit vom System angelegt. Dementsprechend wurden 320 der 386 markierten Kühe automatisch erkannt. Einige Tiere haben die RFID-Marken verloren, wenn sie am Fressgitter hängen geblieben sind, andere mussten den Betrieb nach Markierung und vor Erkennung produktionsbedingt verlassen. Wieder andere sind vermutlich nie dicht genug mit der RFID an der Antenne entlanggelaufen.

5.2 Datenverarbeitung

Für die Datenauswertung wurde mit einem Differenzbild gearbeitet, anhand dessen Umgebungselemente wie Boden, Wände und andere feste Gegenstände herausgefiltert werden konnten. Mit dieser Methode arbeiteten auch schon POURSAHERI et al. (2010).

Sie wandelten jedes fünfte Bild der 2D Videos in ein JPEG um und konnten die Position der Kuh auf dem Bild mit Hilfe eines Logarithmus und dem Differenzbild ermitteln. Da die Kamera bei POURSAERI et al. (2010) in 8 m Entfernung stand, waren die Konturen der Kühe auf den Bildern entsprechend ungenau. Eine kürzere Entfernung der Kamera zum Objekt sorgt somit nicht nur für ein schärferes und detaillierteres Bild, sondern auch für eine bessere Be- und Auswertbarkeit. Sowohl im hier beschriebenen Projekt als auch bei POURSAERI et al. (2010) und VIAZZI et al. (2014) wurden die Bilder in Binärbilder umgewandelt.

5.3 Eruierung von Messpunkten

Wie auch bei VIAZZI et al. (2014) und VAN HERTEM et al. (2017) wurde im Dissertationsprojekt die Wirbelsäule als der höchste Bereich in der Mitte des Kuhrückens definiert. Die Wirbelsäule wurde jeweils mit einer Polyfit-Funktion nachgezeichnet.

Zur Identifizierung der Wirbelsäule aus den Tiefenbildern wurde ähnlich wie bei JABBAR et al. (2017) verfahren, die auch den selben 3D Sensor genutzt haben. Es wurden mit einem Algorithmus die Hüfthöcker und die Wirbelsäule der Kuh segmentiert. Im Gegensatz zu den Auswertungen in diesem Projekt, haben JABBAR et al. (2017) dynamische Messungen durchgeführt. Sie haben die mittleren und maximalen Schwankungen von Wirbelsäule und Hüfthöckern berechnet. Dies ist durch die Abstandsmessung des 3D Sensors einfach möglich, wenn in jedem Bild dieselben Punkte identifiziert werden können. Der Anteil der erfolgreich segmentierten Bilder lag bei 85,7 %. Eine Kontrolle der Berechnungspunkte ermöglichte eine Anpassung des Algorithmus, sodass schlussendlich die Erkennungsrate der Hüfthöcker bei 96 % und der Wirbelsäule sogar bei 100 % lag. Diese Überprüfungen konnten in den Untersuchungen dieses Dissertationsprojekts nur an wenigen ausgewählten Bildern durchgeführt werden, wodurch keine konkrete Aussage über die Genauigkeit der korrekten Segmentierung der Wirbelsäule gegeben werden kann.

JABBAR et al. (2017) und HANSEN et al. (2018) haben, im Gegensatz zum Algorithmus in diesem Projekt, die gesamte Wirbelsäule der Kuh nachgezeichnet. Da die Intention des Projekts darin liegt, Lahmheiten frühestmöglich zu erkennen, bei einem Locomotion Score von 2 eine Krümmung jedoch nur leicht über der Schulter im ersten Drittel der Wirbelsäule erkennbar ist, wurde auf die Darstellung und Verfolgung der

Wirbelsäule im hinteren Drittel der Kuh verzichtet. Es wurde davon ausgegangen, dass im letzte Drittel der Wirbelsäule keine für den Locomotion Score relevante Krümmung mehr stattfindet.

Als Startpunkt der Polylinie wurde dementsprechend der Punkt gewählt, an dem die Gerade zwischen den Hüfthöckern die Wirbelsäule schneidet. In den ersten Berechnungen wurde jedoch kein Endpunkt vorgegeben. Da automatisch die Linie weiter nach vorn bis zum Kopf gezeichnet wird, ist davon ausgegangen worden, dass die Haltung von Kopf und Hals die Berechnungen beeinflusst. Da der Hals der Kuh sehr flexibel ist und die Rückenlinie erheblich verlängert, wurde angenommen, dass die Kreisapproximation bei einer hohen Halshaltung kleinere Krümmungen im signifikanten Rückenbereich über der Schulter ausgleicht und so zu einem besseren Ergebnis führt bzw. dass bei tiefer Halshaltung die Linie automatisch stärker gekrümmt ist und ein schlechterer Score berechnet wird. Deshalb wurde in Berechnung 2 ein Endpunkt der Polylinie vorgegeben. Dazu wurde auf dem Bild des 3D Sensors die Breite der Kuh, ausgehend vom linken Bildrand, gemessen. Da die Kuh an der Schulter anatomisch bedingt breiter wird, wurde ein Algorithmus entwickelt, der das Ende der Polylinie an dem Punkt definiert, an dem der Kuhkörper sich verbreitert und die Schulterpartie beginnt. Auch VIAZZI et al. (2014) und VAN HERTEM et al. (2017) haben den Kopf der Kuh ausgeblendet, um diesen nicht in die Bewertung einfließen zu lassen. Eine genaue Vorgehensweise wird nicht beschrieben. POURSABERI et al. (2010) haben die Abtrennung von Hals und Kopf in ähnlicher Weise wie in diesem Projekt vorgenommen. Es wurde ein Grenzwert für die Summe der weißen Pixel auf der x-Achse des Bildes festgelegt. Das Prinzip entspricht der Berechnung der Breite des Kuhkörpers.

Zur Berechnung des Locomotion Scores wurde im nächsten Schritt ein Kreis an die definierte Polylinie approximiert. Nach einer ähnlichen Methode sind auch POURSABERI et al. (2010) vorgegangen, nur dass sie nicht mit einem 3D Sensor gearbeitet haben, sondern diese Kreisapproximation an den JPEG's aus den 2D Videos vorgenommen haben. Der Durchschnitt der berechneten Radien der Kreise wurde der jeweiligen Kuh zugeordnet. Es wurden jedoch immer nur 4 Bilder ausgewertet und zwar je zwei auf denen das rechte bzw. das linke Hinterbein den Boden berührt. In der hier vorliegenden Arbeit wurden 1 - 60 Bilder ausgewertet und statt des Mittelwertes der Median berechnet, um mögliche Ausreißer, welche z.B. durch nicht korrekt berechnete

Polylinien entstehen können, auszublenden. Da Ausreißer bei vier Bildern eine deutlich stärkere Auswirkung haben, ist die in dieser Arbeit beschriebene Methode - möglichst viele Bilder zur Auswertung zu nutzen - vorzuziehen.

VIAZZI et al. (2014) und VAN HERTEM et al. (2017) haben an die Polylinien keinen Kreis angelegt, sondern den höchsten Punkt berechnet und anhand von Ellipsen und Winkeln von diesem Punkt ausgehend eine Berechnung durchgeführt, die am Ende auch zu einer Aussage über den Krümmungsgrad des Rückens führt. Hier wurde ebenso wie bei POURSAHERI et al. (2010) der Durchschnitt der einzelnen Parameter zur Berechnung des Locomotion Scores herangezogen. Allerdings standen durchschnittlich nur 2,5 Bilder pro Kuh zur Verfügung. Im Gegensatz zu den bisher in der Literatur beschriebenen Verfahren wurde somit im vorliegenden Projekt die größte Anzahl an Bildern zur Berechnung des Locomotion Scores genutzt, wodurch eine höhere Genauigkeit in der Auswertung entsteht.

5.4 Visuelles Locomotion Scoring

Das visuelle Locomotion Scoring wurde überwiegend von Beobachter 1 durchgeführt. Von 320 Kühen, von denen Videodaten zur Auswertung zur Verfügung standen, hat dieser 200 Kühe und 1489 Datensätze bewertet. Bei den in der Literatur beschriebenen Untersuchungen wurden meist deutlich weniger Tiere in die Versuche eingeschlossen. FLOWER et al. (2005), BAHR et al. (2008), JABBAR et al. (2017) und HANSEN et al. (2018) hatten jeweils weniger als 40 Tiere zur Auswertung. Auch die Testzeiträume waren sehr unterschiedlich. FLOWER et al. (2005) haben Messungen an sieben aufeinanderfolgenden Tagen durchgeführt, POURSAHERI et al. (2010) einmal pro Woche an drei aufeinanderfolgenden Wochen und JABBAR et al. (2017) hatten mindestens drei Scores pro Kuh aus einem Zeitraum von drei Wochen. Nur die Untersuchungen von VAN HERTEM et al. (2017) wurden über einen längeren Zeitraum durchgeführt. Hier standen Daten von einem Jahr zur Verfügung. Der Testzeitraum der vorliegenden Arbeit betrug fünf Monate. Gegenüber den meisten Untersuchungen ist der Zeitraum, in welchem das System getestet werden konnte, deutlich länger und die Daten dementsprechend besonders aussagekräftig. Das System konnte beweisen, dass es über einen langen Zeitraum unter Praxisbedingungen funktioniert sowie dauerhaft und störungsfrei Daten aufzeichnen kann. Diese Eigenschaft hebt es deutlich von den

anderen Systemen ab, die jeweils nur für einen kurzen Zeitraum in der Praxis getestet wurden.

Die Anzahl der Datensätze variierte in der vorliegenden Arbeit für die einzelnen Auswertungen. Die größte Stichprobe umfasste 200 Kühe mit 1489 Datensätzen. Da die Anzahl der Datensätze der Anzahl der Videos entsprach, war die Stichprobe in diesem Projekt deutlich größer als in der Literatur bisher beschrieben. JABBAR et al. (2017) hatten 70 Datensätze, POURSAHERI et al. (2010) 156, SCHLAGETER-TELLO et al. (2018) 223, FLOWER et al. (2005) 266, PLUK et al. (2012) 300 und RAJKONDAWAR et al. (2006) 700. Durch die deutlich erhöhte Anzahl an Datensätzen ist die Stichprobe demnach repräsentativer als die meisten bisher durchgeführten Untersuchungen. Doch die Untersuchungen von VAN HERTEM et al. (2017) können eine größere Stichprobe als dieses Projekt vorweisen, denn auch sie hatten die Prüfung der Praxistauglichkeit zum Ziel.

Bei den visuellen Beurteilungen wurden Besonderheiten, die während der Bonitur auftraten, für den jeweiligen Datensatz vermerkt. Solche Besonderheiten waren beispielsweise „Darmmentleerung“, „Stau“ oder „verkoteter Rücken“. Diese besonderen Vorkommnisse, die während der Bonitur beobachtet wurden, können einen Einfluss auf den vom 3D Sensor berechneten Locomotion Score haben, weil beispielsweise die Kuh den Rücken naturgemäß bei der Darmmentleerung vermehrt krümmt. Auch PASTELL et al. (2008) sind solche Datensätze negativ aufgefallen und wurden gelöscht. Die potentiell fehlerhaften Datensätze dieses Dissertationsprojekts wurden in der Berechnung mit Stichprobe 4 (siehe Tabelle 4) herausgefiltert, um verfälschte Scores nicht in die Auswertung einzubeziehen. Übrig blieben 1052 Datensätze von 181 Kühen. Die Tatsache, dass bei 400 Datensätzen (29,3 %) ein besonderes Vorkommnis vermerkt wurde, welches vom 3D Sensor noch nicht automatisch erkannt werden kann, zeigt, dass der Einsatz in der Praxis noch deutlich mehr verlangt als die bloße Erkennung der Rückenlinie und Berechnung eines Scores. Um in der Praxis zu funktionieren, muss ein System zusätzlich bestimmte Abnormitäten erkennen, die dann die Verhinderung der Ausgabe eines Scores zur Folge haben müssen. Innerhalb der Projektlaufzeit war die Entwicklung eines solchen Tools nicht mehr möglich, sollte aber bei Anschlussprojekten unbedingt berücksichtigt werden.

Die Validierung durch die visuelle Bonitur erfolgte in den Untersuchungen anderer Wissenschaftler teilweise durch Veterinäre (RAJKONDAWAR et al. 2006, PASTELL et al.

2008, BICALHO et al. 2007, VIAZZI et al. 2014), teilweise durch wissenschaftliche Beobachter (BAHR et al. 2008, SONG et al. 2008, PLUK et al. 2012, JABBAR et al. 2017, VAN HERTEM et al. 2017, SCHLAGETER-TELLO et al. 2018, HANSEN et al. 2018). Im Dissertationsprojekt wurden die meisten Bewertungen durch eine erfahrende wissenschaftliche Beobachterin durchgeführt. Es wäre sicher von Vorteil gewesen, einen Veterinär vor Ort zu haben, der wie bei RAJKONDAWAR et al. (2006) und PASTELL UND KUJALA (2007) auch direkt Läsionen an den Klauen beurteilt, doch in den meisten Untersuchungen werden nur die Bewegungsmuster beurteilt.

Das Locomotion Scoring ist in vielen Versuchen nach der Methode von SPRECHER et al. (1997) durchgeführt worden (PASTELL UND KUJALA 2007, PASTELL et al. 2008, VAN HERTEM et al. 2014, JABBAR et al. 2017, HANSEN et al. 2018). Einige Wissenschaftler haben auch Modifikationen vorgenommen (RAJKONDAWAR et al. 2006, BAHR et al. 2008) bzw. die Definition von „gesund“ und „lahm“ für sich angepasst. Häufig werden Locomotion Scores (LS) von 1 und 2 als gesund bewertet (PASTELL UND KUJALA 2007, ITO et al. 2010, VIAZZI et al. 2014, VAN HERTEM et al. 2017, SCHLAGETER-TELLO et al. 2018, HANSEN et al. 2018) und Scores von 3 oder schlechter als lahm. Eine solche Milderung der Bewertung wurde in der vorliegenden Arbeit nicht vorgenommen. Hier wurden nur Kühe mit einem LS von 1 als gesund bewertet, denn das Ziel des Projektes ist, Lahmheiten frühestmöglich zu erkennen und nicht die ersten Symptome der Lahmheit zu ignorieren. Für ein automatisches System ist es jedoch leichter, ein fortgeschrittenes Stadium der Lahmheit zu erkennen als ein beginnendes. Vermutlich wurden die Anpassungen in der Bewertung vorgenommen, um die Ergebnisse der Untersuchungen eindeutiger zu machen.

Um die Objektivität des visuellen LS zu erhöhen, können auch mehrere Beobachter eingesetzt werden. VAN HERTEM et al. (2017) und SCHLAGETER-TELLO et al. (2018) hatten zwei Beobachter, BAHR et al. (2008) und HANSEN et al. (2018) hatten drei Beobachter, ebenso wie BICALHO et al. (2007) in ihrer zweiten Studie und SONG et al. (2008) hatten sogar vier subjektive Bewerter. Im Dissertationsprojekt wurde auch eine kleine Stichprobe durch drei Beobachter bewertet, um die Aussagefähigkeit und Objektivität zu erhöhen. Insgesamt wurden 40 Datensätze von 40 Kühen von dem Beobachter-Trio bonitiert. Dabei wurde wie bei BICALHO et al. (2007) der Mittelwert der drei Boniturnoten genommen und auf eine ganze Zahl gerundet. Die Anzahl der Tiere ist vergleichbar mit denen der Untersuchungen von FLOWER et al. (2005), BAHR et

al. (2008) und JABBAR et al. (2017), die Anzahl der Datensätze jedoch deutlich geringer, da je Kuh nur eine Bewertung vorgenommen wurde. Das Trio bestand aus zwei wissenschaftlich geschulten Beobachterinnen und einem unerfahrenen Beobachter. Die Bonitur erfolgte unabhängig voneinander anhand der Videoaufnahmen. So konnten sich die Beobachter nicht gegenseitig beeinflussen. Die Untersuchungen von BRENNINKMEYER et al. (2007) haben gezeigt, dass die Übereinstimmung zwischen den Beobachtern mit der Anzahl der Durchführungen steigt. Von einem solchen Effekt kann in der vorliegenden Studie nicht ausgegangen werden, da dafür die Anzahl der Bonituren zu gering ist. Bei BRENNINKMEYER et al. (2007) zeigte sich auch eine zufriedenstellendere Übereinstimmung bei untrainierten Beobachtern, wenn die Skala der Bewertungsmöglichkeiten von vier auf zwei minimiert ist. Das Beobachter-Trio dieser Studie hatte jedoch fünf Scores zur Bonitur zur Verfügung. Schlussendlich ist die Aussagekraft der Boniturnoten und der Grad der Objektivität des Beobachter-Trio dennoch deutlich höher als bei nur einem Beobachter.

Ob die Bonitur der Kühe live oder anhand von Videomaterial erfolgte, ist bei den in der Literatur beschriebenen Untersuchungen sehr unterschiedlich. Teilweise wurden auch beide Beurteilungsmethoden parallel in einer Untersuchung angewandt (JABBAR et al. 2017, SONG et al. 2008). Da sich in der ersten Versuchsphase des Dissertationsprojekts die live Bonitur als nicht für den Dauerversuch geeignet herausstellte, weil die Kühe von der Person zu sehr abgelenkt waren, kam eine live Bonitur nicht in Frage. Eine Bonitur wie bei JABBAR et al. (2017) im Laufstall, nachdem die Kühe die Versuchsumgebung verlassen haben, wäre zwar möglich gewesen, jedoch sollte die Bewertung des Gangs von Mensch und Maschine möglichst exakt im selben Moment geschehen, um besser dasselbe Bild darzustellen. Läuft die Kuh zum Beispiel im Laufstall auf einem anderen Untergrund als in der Versuchsumgebung, kann dieser die Boniturnote bereits beeinflussen.

Validierung und Klassifizierung der automatischen Scores

Mit diesem Dissertationsprojekt gelang es erstmalig in Deutschland einen 3D Sensor zur Lahmheitsbeurteilung in der Praxis zu etablieren und dieses System anhand visueller Tierbeobachtungen zu validieren. Ein besonders positiver Effekt in dem Projekt war die Entwicklung der Algorithmen während der praktischen Versuchsphase.

Dadurch wurde kontinuierlich mehr Versuchsmaterial (Bildmaterial) gewonnen und bei jeder Anpassung der Algorithmen, für eine verbesserte Berechnung der automatischen Scores, konnte eine größere Stichprobe herangezogen werden. Insgesamt erfolgten drei Anpassungen bzw. Auswertungen.

5.5 Auswertung 1

Die erste Auswertung erfolgte anhand der vorgegebenen Klassifizierung der Krümmungsgrade in Tabelle 3. Diese ergibt eine Übereinstimmung der aLS mit den vLS von 50 %. Da dieser Wert nicht zufriedenstellend ist, wurde eine Minimierung der Klassen vorgenommen, ähnlich der Studien von BAHR et al. (2008) und PLUK et al. (2012). Statt fünf Scores gab es nur noch drei Scores, wobei alle Scores ≥ 3 als 3 ausgegeben wurden. Dadurch kann eine Übereinstimmung von 64 % erreicht werden. Auf diese Methode wurde zurückgegriffen, da eine Abstufung in gesund, leicht lahm und lahm ausreichend ist, weil es den Landwirt weniger interessiert, wie stark lahm eine Kuh ist, sondern ob sie a) gesund ist und somit keine Behandlung notwendig ist, ob sie b) leicht lahm ist und einer Lahmheit durch Untersuchung und Behandlung entgegengewirkt werden kann oder ob das Tier c) lahm ist und behandelt werden muss. Da der Anspruch des Projekts, auch mit dem automatischen System eine 5-stufige Klassifizierung ausgeben zu können, nicht von vornherein abgelehnt werden sollte, wurden in allen weiteren Auswertungen die Übereinstimmungen sowohl für eine fünfstufige Klassifizierung als auch für eine dreistufige Klassifizierung ausgegeben. Auch VAN HERTEM et al. (2014) haben versucht, eine 5-stufige Bewertung zu erreichen. Doch auch hier lag die Übereinstimmung maximal bei 53 %. Sie sind schlussendlich auf eine zweistufige Bewertung zurück gegangen, um aussagekräftige Ergebnisse zu erreichen. Auch VIAZZI et al. (2014) haben nur eine zweistufige Bewertung vorgenommen.

5.6 Auswertung 2

Da auch die 64 % Übereinstimmung bei der 3-stufigen Klassifizierung nicht zufriedenstellend sind, wurde der Algorithmus für die Berechnung des aLS angepasst. Durch das definierte Ende der Polylinie soll die Kopfhaltung weniger Einfluss auf die berechnete Rückenkrümmung haben. Außerdem wurde eine kleinere Stichprobe gewählt. Da die visuelle Bewertung stark subjektiv ist, wurde für die zweite

Berechnung nur noch die jeweils erste Bewertung von jeder Kuh ausgewertet. Dadurch verringerte sich die Stichprobe auf 197 Datensätze, aber die Subjektivität beim vLS wird verringert. Da bei der Beurteilung der vLS immer alle Datensätze einer Kuh hintereinander geschaut wurden, könnte sich Beobachter 1 durch seine vorangegangene Bewertung selbst beeinflusst haben. Werden jedoch nur die jeweils ersten Datensätze einer Kuh ausgewertet, kann davon ausgegangen werden, dass diese Daten weniger subjektiv sind. Schlussendlich wurden außerdem nur noch die Bilddaten von Sensor 1 zur Berechnung der aLS herangezogen, weil auch hier davon ausgegangen wird, dass die Daten von Sensor 2 teilweise die Daten verfälschen, weil beispielsweise bei Stau im Versuchsgang auf diesem Sensor eine andere Kuh abgebildet sein kann als auf Sensor 1. Dadurch wurde zwar die Berechnung gegenüber anderer Versuche (VAN HERTEM et al. 2014, JABBAR et al. 2017) erschwert, da weniger Schritte abgebildet werden, aber es konnte ein Störfaktor eliminiert werden.

Der direkte Vergleich der aLS und vLS zeigt für Berechnung 2 mit Stichprobe 2 eine Übereinstimmung von 56 % und ist damit bereits höher als bei der 5-stufigen Bewertung von VAN HERTEM et al. (2014). Überraschenderweise hat sich dagegen die Übereinstimmung der 3-stufigen Bewertung zur vorherigen Berechnung verschlechtert. Hier werden nur noch 58 % erreicht. Da von Berechnung 1 zu Berechnung 2 mehrere Faktoren verändert wurden (Größe der Stichprobe, Bilder ausschließlich von Sensor 1, definiertes Ende der Polylinie), wurde der Algorithmus von Berechnung 1 auch auf Stichprobe 2 angewandt, ausschließlich von den Daten des Sensors 1. So kann ein direkter Vergleich zwischen dem Algorithmus von Berechnung 1 und Berechnung 2 erfolgen. Es zeigt sich, dass die Übereinstimmung sowohl bei der 5- als auch bei der 3-stufigen Beurteilung bei Berechnung 1 leicht unter der von Berechnung 2 liegt (Tabelle 6, Tabelle 7). Es kann somit davon ausgegangen werden, dass eine Begrenzung der Polylinie zu einer genaueren Auswertung führt. Wie schon von VIAZZI et al. (2014) beschrieben, sollten Kopf und Hals der Kuh nicht mit ausgewertet werden.

Der Vergleich der Häufigkeiten der einzelnen aLS der Berechnungen mit und ohne Hals sowie der vLS (Abbildung 20) zeigt, dass visuell deutlich häufiger der Score 1 vergeben wurde als vom automatischen System. Deshalb wird vermutet, dass die Klassengrenze für den Score 1 im automatischen System zu eng ist. Im Gegensatz dazu liegen die Häufigkeiten der visuellen Bonitur der Scores 4 und 5 unter denen des automatischen Systems, sodass die Grenzen zu hoch scheinen.

Tabelle 9 stellt in einer Matrix jeweils die Anzahl der Abweichungen der aLS von den vLS je Größe der Abweichung und je visuellem Score für den Algorithmus mit und ohne Hals dar. Die Matrix „mit Hals“ lässt keine Trends oder Rückschlüsse zu. Die Matrix „ohne Hals“ zeigt dagegen deutlich, dass zumeist der aLS mindestens eine Note unter dem vLS liegt (Tabelle 9). Diese Erkenntnis untermauert noch einmal die Entscheidung für eine automatische Beurteilung ausschließlich des Kuhrückens ohne Hals und zeigt, dass eine Optimierung des Algorithmus durch iterative Anpassung der anfangs gesetzten Grenzen für die einzelnen Scores notwendig ist. Das Ergebnis der als optimal berechneten Klassengrenzen (Tabelle 10) zeigt, dass eine Einteilung der Klassen mit einer Kommastelle nicht ausreichend ist. Dementsprechend wurden die Grenzen mit zwei Kommastellen ausgegeben. Das zeigt wiederum, wie empfindlich und genau die Messungen des automatischen Systems sind. Eine solche Anpassung der Klassengrenzen, die möglichst viele Kühe richtig klassifiziert, haben auch VAN HERTEM et al. (2014) vorgenommen. Sie erreichten dadurch eine Übereinstimmung von 53 %.

5.7 Auswertung 3

5.7.1 Stichprobe 3

Der Algorithmus mit den neu definierten Klassengrenzen wurde an Stichprobe 3 überprüft (Tabelle 11). Das Ergebnis einer Übereinstimmung von 54 % bzw. 63 % zeigt eine deutliche Verbesserung bei dem 3-stufigen System und eine geringe Verschlechterung des 5-stufigen Systems gegenüber der vorangegangenen Auswertung. Die geringfügige Verschlechterung im 5-stufigen Modell ist durch die größere Stichprobe zu erklären und die Verbesserung im 3-stufigen System zeigt, dass die neuen Klassengrenzen besser mit den vLS übereinstimmen. Es kann davon ausgegangen werden, dass ein Großteil der in Auswertung 2 einen Score zu tief eingeordneten Tiere nun im richtigen Score ist. Zumindest trifft dies bei den Tieren mit einem Score von 1 und 2 zu, denn bei der 3-stufigen Einteilung ist eine Fehlzurordnung in einen tieferen Score ab der Note 3 irrelevant. Das 5-stufige Modell erreicht mit einer Übereinstimmung von 54 % sogar eine bessere Übereinstimmung als VAN HERTEM et al. (2014).

Abbildung 21 veranschaulicht, welche Messwerte das automatische System berechnet hat und welche visuellen Scores entsprechend vergeben wurden. Diese Darstellung zeigt, dass die Messwerte des automatischen Systems in den einzelnen visuellen Scores

sehr stark streuen. Dadurch unterscheiden sich die Messwerte im vLS 1 und 2 kaum. Diese Unterscheidung ist jedoch die Differenzierung in „gesund“ und „leicht lahm“, was wiederum ein Ziel dieser Arbeit ist. Es zeichnet sich demnach ab, dass eine automatische Unterscheidung zwischen „gesund“ und „leicht lahm“ sehr schwer ist. Dies erklärt, warum VIAZZI et al. (2014) und HANSEN et al. (2018) eine Einteilung in „LS 1 und 2 = gesund“ und „LS > 3 = lahm“ vorgenommen haben.

Auffällig in Abbildung 21 sind aber auch die Ausreißerwerte > 1,2. Entsprechend wurde eine neue Stichprobe erstellt, bei welcher Datensätze mit einem Messwert > 1,2 automatisch ausgeschlossen sind. Die Vermutung liegt nahe, dass einige dieser Datensätze mit Messfehlern aus den Bildern entstanden sind, bei denen im visuellen Locomotion Scoring eine Besonderheit vermerkt wurde. Entsprechend wurden solche Datensätze aus der neuen Stichprobe ausgeschlossen. Eine solche manuelle Löschung von Datensätzen, die eine erschwerte Auswertung nach sich ziehen, haben auch PASTELL et al. (2008) vorgenommen. Zusätzlich wurden Datensätze, bei denen der automatische und der visuelle Score mehr als eine Note voneinander abweichen, aus der Auswertung ausgeschlossen. Insgesamt waren dies 56 Datensätze, also 3,76 % aller Datensätze von Stichprobe 3.

5.7.2 Stichprobe 4

Tabelle 12 zeigt die Ergebnisse des Vergleichs von vLS und aLS nach den zuvor beschriebenen Anpassungen. Im 5-stufigen Modell wird eine Übereinstimmung von 59 % und im 3-stufigen von 65 % erreicht. Eine höhere Übereinstimmung kann vermutlich nur erreicht werden, wenn - wie bei POURSAHERI et al. (2010) und VAN HERTEM et al. (2014) beschrieben - kuhindividuelle Klassengrenzen verwendet werden. Werden die Messwerte der Rückenkrümmung mit Hilfe eines Box-Plots in den dazugehörigen vLS dargestellt, zeigen sich deutlich die einzelnen Klassen der Scores (Abbildung 22). Das bestätigt noch einmal die vorgenommene Optimierung der Klassengrenzen.

Die statistische Überprüfung von Berechnung 3 mit Stichprobe 4 hat nach Spearman eine Korrelation von 0,73 ergeben (Tabelle 13). Das bedeutet, dass zwischen der Krümmung des Rückens (Messwerte Rückenkrümmung) und dem visuellen Locomotion Score eine enge Verbindung besteht. Auch die Varianzanalyse ANOVA bestätigt einen signifikanten Einfluss des vLS auf den Messwert der Rückenkrümmung.

Dadurch ist statistisch nachgewiesen, dass sich das entwickelte System zur automatischen Erfassung des Locomotion Scores eignet. Die Unterscheidung der einzelnen Scores voneinander wird anhand des Tukey HSD Tests überprüft, da das Box-Plot (Abbildung 22) eine signifikante Unterscheidung nicht eindeutig darstellt. Der Test zeigt jedoch, dass sich die einzelnen Klassen hoch signifikant voneinander unterscheiden.

In der Literatur wird häufig die Sensitivität und die Selektivität eines neuen Systems berechnet. Der Algorithmus der Berechnung 3 ergab an Stichprobe 4 eine Sensitivität von 96 % und eine Spezifität von 48 %. Verglichen mit dem StepMetrixTM-System, welches in der Studie von RAJKONDAWAR et al. (2006) nur eine Sensitivität von 68 % erreicht, ist die automatische Erkennung lahmer Kühe deutlich sicherer. MAERTENS et al. (2011) haben das System weiterentwickelt und insgesamt 20 kinematische Variablen gemessen. Dadurch erreichen sie zwar eine Sensitivität ihres Systems von 76 - 90 %, können jedoch nicht die 96 % des vorliegenden Projekts erzielen. Im Gegensatz dazu liegt die Spezifität bei RAJKONDAWAR et al. (2006) bei 85 %, bei MAERTENS et al. (2011) sogar bei 86 - 100 % und in der vorliegenden Untersuchung bei nur 48 %. Jedoch ist das Ziel dieses Projekts die Erkennung kranker Kühe und nicht die Erkennung gesunder Kühe. Somit ist das in dieser Arbeit beschriebene System dem StepMetrixTM-System überlegen. Eine Beurteilung per Bilddaten ist demnach einer Bewertung per Auftrittskraftmessung vorzuziehen.

Auch gegenüber indirekten Messungen, wie beispielsweise der Milchmenge und Kraftfutteraufnahme pro Tag in den Untersuchungen von DE MOL et al. (2013) (86 % Sensitivität), liegt die Sensitivität dieses Projekts um 10 %-Punkte höher.

Die Sensitivität von Bildverarbeitungstechnik kann aber auch deutlich höher sein. POURSAHERI et al. (2010), welche ebenfalls mit einer Kreisapproximation an der Rückenlinie der Kühe gearbeitet haben, erreichen eine Sensitivität von 100 % und eine Spezifität von 98 %. Diese Ergebnisse zeigen, dass sich die Kreisapproximation zur Auswertung der Rückenkrümmung grundsätzlich sehr gut eignet. Eine Markteinführung des Systems erfolgte jedoch nicht; vermutlich, weil die Videoaufnahmen von der Seite unter Praxisbedingungen schwer umzusetzen sind. Die Anforderungen an die Systemumgebung und der Platzbedarf mit freier Sicht sind in kaum einem Stall zu finden.

Die Untersuchungen von VIAZZI et al. (2014), welche als erste einen 3D Sensor zur Lahmheitsbeurteilung genutzt haben, zeigen eine Sensitivität von 82 % und eine Spezifität von 91 %. Das bedeutet, auch in dieser Studie konnten die gesunden Tiere sicherer erkannt werden als die lahmen. Das vorliegende System erkennt kranke Kühe um 14 %-Punkte besser als das von VIAZZI et al. (2014). SCHLAGETER-TELLO et al. (2018) haben unter dem Versuchsaufbau von VAN HERTEM et al. (2017) nur eine Sensitivität von 74 % erreicht und eine Spezifität von 68 %. Jedoch standen bei diesen Untersuchungen auch nur zwei Ergebnisse pro Kuh zur Verfügung. JABBAR et al. (2017) entwickelten das System weiter und erreichten schlussendlich eine Sensitivität von 100 % und eine Spezifität von 75 %. Doch ihre Untersuchung umfasste lediglich eine Stichprobe von 22 Kühen. Von diesen waren 19 Tiere lahm (LS 2-5) und drei Tiere gesund (LS 1). Diese Stichprobe spiegelt keine realen Bedingungen in einem Milchviehstall wider. Verglichen mit den 1052 Tieren in diesem Projekt ist die Stichprobe nicht aussagekräftig. Diese Gegenüberstellung verdeutlicht eher das besonders gute Ergebnis von 96 % in diesem Projekt.

5.7.3 Stichprobe 5

Stichprobe 5 entspricht der Stichprobe 4, aber von jeder Kuh wurde nur die erste Bewertung genutzt. Auf diese Weise soll die Subjektivität der Bewertung weiter reduziert werden. In den Auswertungen wurden immer alle Datensätze einer Kuh hintereinander angesehen. Es wird davon ausgegangen, dass sich Beobachter 1 durch die zuvor vergebene Note unterbewusst selbst beeinflusst haben könnte. Eine durchmischte Reihenfolge der ausgewerteten Bilder wäre für zukünftige Auswertungen ratsam. Der Vergleich der aLS und vLS zeigt bei fünf Stufen eine Übereinstimmung von 56 % und bei drei Stufen von 64 % (Tabelle 15). Im direkten Vergleich mit den Ergebnissen von Stichprobe 4, welche die Ergebnisse von Stichprobe 5 quasi beinhalten, zeigt sich, dass sich im 5-stufigen Modell die Übereinstimmung nur minimal verändert hat. Nur 3 %-Punkte weniger Übereinstimmungen konnten festgestellt werden. Das deutet darauf hin, dass sowohl Beobachter 1 homogen bewertet als auch die Messergebnisse der Rückenkrümmung gleichmäßig und gut sind.

Das Box-Plot-Diagramm (Abbildung 23) zeigt zum einen die deutlich erkennbaren Klassen der einzelnen Scores, zum anderen wird aber auch ersichtlich, dass sich die Messwerte in Bezug auf die visuellen Scores überschneiden. Besonders auffällig ist dies

beim Vergleich der Messwerte in vLS 4 und 5. Das Box-Plot in vLS 5 wird vollständig vom Box-Plot in vLS 4 abgedeckt. Daraus ergibt sich die Annahme, dass sich die vLS 4 und 5 nicht signifikant voneinander unterscheiden. Diese Hypothese wurde statistisch überprüft (Tabelle 16). Es zeigt sich eine starke Korrelation zwischen Rückenkrümmung und visuellem Locomotion Score ($r = 0,70$). Verglichen mit der Korrelation von Stichprobe 4 ist diese in Stichprobe 5 minimal geringer. Ebenso hat sich auch die Übereinstimmung leicht verringert. Das widerlegt die Vermutung, dass Beobachter 1 sich selbst beeinflusst hat. In diesem Fall hätte eine Verbesserung statt einer Verschlechterung der Übereinstimmung und Korrelation eintreten müssen.

Der Einfluss des vLS auf die automatisch generierten Messwerte der Rückenkrümmung ist signifikant ($p < 0,05$). Der Tukey HSD Test bestätigt die Annahme, dass sich die Messwerte in den visuellen Scores 4 und 5 nicht unterscheiden ($p = 0,76$). Alle anderen Scores unterscheiden sich dagegen hoch signifikant ($p < 0,05$). Dieses Erkenntnis spricht dafür, dass von der 5-stufigen Beurteilung abgewichen und eine vereinfachte Bewertung erfolgen sollte. In der Literatur wird teilweise eine modifizierte Bewertung vorgenommen, die nur drei Scores enthält (AMORY et al. 2004, BAHR et al. 2008, PLUK et al. 2012). Eine solche Bewertung scheint aufgrund der Erkenntnis, dass die Unterscheidung zwischen Score 4 und 5 nicht signifikant ist, sinnvoll. Zumindest eine Reduktion auf vier Scores ist ratsam, wie bereits von COOK (2003) beschrieben.

Es stellt sich die Frage, ob die Übereinstimmung des visuellen und des automatischen Locomotion Scores innerhalb der einzelnen vLS unterschiedlich ist. Tabelle 17 zeigt, dass der vLS 2 besonders gut vom automatischen System erkannt wird (Übereinstimmung = 80 %). Mit diesem Ergebnis wurde die Zielstellung dieser Arbeit - das automatische Erkennen einer beginnenden Lahmheit - erreicht. Es ist jedoch nicht zu missachten, dass die Übereinstimmung der anderen Scores unter 50 % liegt, abgesehen von vLS 5, welcher aber vernachlässigt werden kann, da der Anteil von 3 Tieren nicht repräsentativ ist.

Bei Stichprobe 5 erreicht das automatische System eine Sensitivität von 95 %. Das ist als sehr gut zu bewerten und sicher auf die hohe Übereinstimmung in Locomotion Score 2 zurückzuführen. Die Spezifität erreicht dagegen lediglich 39 %.

Verglichen mit den Ergebnissen von RAJKONDAWAR et al. (2006), MAERTENS et al. (2011), VIAZZI et al. (2014) und SCHLAGETER-TELLO et al. (2018), welche keine höhere Sensitivität als 90 % erreichen, arbeitet dieses System deutlich besser und ist damit

praxistauglicher. Nur JABBAR et al. (2017) haben eine höhere Sensitivität (100 %), aber auch eine deutlich kleinere Stichprobe (23 Kühe). Auch VAN HERTEM et al. (2014) erreichen nur eine Sensitivität von 55 %, dafür aber eine Spezifität von über 90 % und eine Genauigkeit von 81 %. In Stichprobe 5 wird in diesem Projekt eine Genauigkeit von 79 % erreicht, also trotz der deutlich geringeren Spezifität ein sehr gutes Ergebnis. Die Genauigkeit von SCHLAGETER-TELLO et al. (2018) liegt wiederum nur bei 69 %, was eventuell auf die geringen Wiederholungen zurückzuführen ist. HANSEN et al. (2018) erreichen mit 83 % die höchste Genauigkeit. Allerdings wurden hier keine Angaben zur Sensitivität und Spezifität gemacht, sodass ein direkter Vergleich schwierig ist.

5.7.4 Stichprobe 6

In der letzten Analyse erfolgte die visuelle Bewertung statt von einer Beobachterin von drei Beobachtern. Durch diese Maßnahme soll die Subjektivität in der visuellen Bewertung noch weiter reduziert werden. In der Auswertung wurde der Mittelwert der drei visuellen Scores (auf eine ganze Zahl gerundet) dem automatischen Score gegenübergestellt, entsprechend dem Vorgehen von BICALHO et al. (2007) in ihrer Studie.

In der direkten Gegenüberstellung der Scores wird eine Übereinstimmung von 67 % im 5-stufigen System erreicht und 72 % im 3-stufigen. Wobei anzumerken ist, dass von den 40 Kühen keine Kuh vom Beobachter-Trio mit der Note 4 oder 5 bewertet wurde. Die schlechteste Note war die 3. Auch das automatische System hat keine 5 vergeben, aber zumindest zweimal die Note 4. Dementsprechend erscheint für diese Auswertung der Stichprobe 6 die 3-stufige Klassifizierung am sinnvollsten und die 72 % Übereinstimmung sind als sehr gutes Ergebnis zu bewerten.

In Abbildung 24 sind die Messwerte zu den einzelnen Scores dargestellt. Für eine bessere Übersicht sind hier die ungerundeten vLS dargestellt. Werden hier nur die Scores betrachtet, bei denen alle Beobachter gleich entschieden haben, also Score 1,0, 2,0 und 3,0, fällt auf, dass die Grenze zwischen vLS 1 und 2 relativ eindeutig kurz unterhalb 0,5 liegt. Wird dieser Grenzwert aber mit den berechneten Grenzwerten in Tabelle 10 verglichen, zeigt sich, dass bei 0,48 bereits die Grenze zwischen aLS 2 und 3 ist. Es stellt sich die Frage, ob gerade in den Fällen, in denen eine eindeutige (einstimmige) Zuordnung zu einem Score visuell schwerfällt, die Messwerte genauer

und eindeutiger sind. In Folgeuntersuchungen sollten Kühe mit diesen „uneindeutigen visuellen Scores“ direkt im Anschluss veterinärmedizinisch untersucht werden, um eine fundiertere Zuordnung machen zu können. Darauf basierend wäre eine neue Einteilung der Klassen für das automatische System denkbar.

An dieser Stelle fällt auch auf, dass das Intervall zwischen den Score-Grenzwerten von 0,29 in LS 2 am größten ist (Tabelle 10). LS 3 hat dagegen nur ein Intervall von 0,16 und LS 4 von 0,13. Die schmalere Grenzen in den höheren Scores könnten ein Indiz dafür sein, dass eine 3-stufige Bewertung sinnvoller ist als eine 5-stufige.

Nichtsdestotrotz kann ein enger Zusammenhang zwischen den Messwerten und den Boniturnoten nachgewiesen werden ($r = 0,73$). Auch bei der 3-stufigen Bewertung unterscheiden sich die einzelnen Scores signifikant voneinander, wobei die Auswertung des Tukey HSD Tests die Annahme von Abbildung 24 bestätigt, dass sich Score 1 und 2 am wenigsten voneinander unterscheiden.

In Stichprobe 6 wird eine Sensitivität von 100 % erreicht. Damit ist das Ergebnis vergleichbar mit dem Ergebnis der Studie von JABBAR et al. (2017). Aufgrund der größeren Stichprobe in diesem Projekt (40 vs. 23), ist das Ergebnis noch aussagekräftiger. Im Vergleich zur Auswertung von Stichprobe 5 mit 181 Tieren, ist die Genauigkeit bei dieser Stichprobe jedoch von 79 % auf 63 % gesunken.

6. Schlussfolgerungen

Folgende Schlussfolgerungen können aus dem Projekt zur Entwicklung eines Bildverarbeitungssystems zur Überwachung des Locomotion Scoring in Milchviehbetrieben gezogen werden:

Prozessentwicklung

- Die Durchführung des Locomotion Scorings durch einen menschlichen Beobachter ist in modernen Milchviehbetrieben schwer durchführbar, da die Kühe zu sehr auf den Menschen reagieren und so einen flüssigen Produktionsablauf verhindern. Die visuelle Bonitur anhand von 2D Videoaufnahmen hat sich im Versuchszeitraum als praxistauglich erwiesen.
- Im Gegensatz zu anderen Untersuchungen wurde sowohl die 2D Kamera dichter und tiefer montiert als auch der 3D Sensor in einer geringeren Höhe. Durch die Einsatzmöglichkeit in Altgebäuden mit begrenzenden räumlichen Möglichkeiten, wird die Praxistauglichkeit des Systems bestätigt und bietet einen Vorteil gegenüber anderen Systemen. Darüber hinaus ergibt sich durch die kürzere Entfernung der Sensoren zum Objekt ein schärferes und detaillierteres Bild, wodurch die Be- und Auswertbarkeit verbessert wird. Für zukünftige Projekte ist geplant, zwei hintereinander geschaltete 3D Sensoren zu triggern, um Bilddaten von zwei Schrittfolgen auswerten zu können.
- Die Segmentierung der Wirbelsäule zur Berechnung des Locomotion Scores hat sich als wirkungsvoller Messpunkt erwiesen. Zur Überprüfung von Messfehlern und Erhöhung der Genauigkeit ist eine Kontrolle der korrekten Segmentierung in den Einzelbildern notwendig und in nachfolgenden Untersuchungen vorgesehen. Auf diese Weise kann auch überprüft werden, ob zusätzliche Algorithmen zur Vorauswahl der Bilder, wie bei VAN HERTEM et al. (2017) und HANSEN et al. (2018) beschrieben, notwendig sind. Denn die Tatsache, dass bei 29 % der Datensätze vom Beobachter ein besonderes Vorkommnis vermerkt wurde, deuten darauf hin, dass vom System auch diese Abnormitäten erkannt werden müssen, um solche Bilder aus der Auswertung auszuschließen.

Schlussfolgerungen

- Erstmals in Deutschland konnte über einen Zeitraum von fünf Monaten ein 3D Sensor zur automatischen Berechnung des Locomotion Scores in einem Milchviehbetrieb integriert werden.

Validierung und Klassifizierung der Algorithmen zur Ableitung des Locomotion Scores

- Zur Berechnung des Locomotion Scores sollten der Kopf und der Hals der Kuh nicht einbezogen werden. Die Übereinstimmung von vLS und aLS ist bei einer Berechnung ohne Hals größer.
- Im 5-stufigen Modell wird eine Übereinstimmung von 59 % und im 3-stufigen von 65 % erreicht. Eine höhere Übereinstimmung kann vermutlich nur erreicht werden, wenn kuhindividuelle Daten verwendet werden.
- Der Unterschied zwischen Score 4 und 5 ist nicht signifikant. Eine Modifizierung der Bewertung mit Reduktion auf vier Scores ist ratsam.
- Die Erkennungsrate einer Kuh mit Locomotion Score 2 ist besonders hoch (80 %), wodurch das Ziel des Projekts, eine beginnende Lahmheit zu erkennen, erreicht wurde.
- Das automatische System ist in der Lage, das Locomotion Scoring für den Menschen zu übernehmen. Es eignet sich besonders gut zur Erkennung einer beginnenden Lahmheit bzw. leicht lahmender Kühe.

7. Literaturverzeichnis

ADR (2017): Rinderproduktion in Deutschland 2016. Jahresbericht der Arbeitsgemeinschaft Deutscher Rinderzüchter e. V. Ausgabe 2017, 55.

AHDB DAIRY (2017): A new approach to lameness control in dairy cows: treatment of individual lame cows and farm level interventions. Agriculture & Horticulture Development Board, Kenilworth (<https://dairy.ahdb.org.uk/research-development/health-welfare/current-projects/lameness-control-in-dairy-cows/#.WIZHz4WcFhE>, abgerufen am 23.01.2017)

ALSAAOD, M.; RÖMER, C.; KLEINMANN, J.; HENDRIKSEN, K.; ROSE-MEIERHÖFER, S.; PLÜMER, L.; BÜSCHER, W. (2012): Electronic detection of lameness in dairy cows through measuring pedometric activity and laying behavior. Applied Animal Behaviour Science 142, 134 – 141.

AMORY, J.R.; KLOOSTERMAN, P.; BARKER, Z.E.; WRIGHT, J.L.; BLOWEY, R.W.; GREEN, L.E. (2006): Risk factors for reduced locomotion in dairy cattle on nine-teen farms in the Netherlands. J. Dairy Sci. 89, 1509–1515.

ARCHER, S.C.; GREEN, M.J.; HUXLEY, J.N. (2010): Association between milk yield and serial locomotion score assessments in UK dairy cows. J. Dairy Sci. 93, 4045 – 4053.

BACH, A.; DINARÉS, M.; DEVANT, M.; CARRÉ, X. (2007): Association between lameness and production, feeding and milking attendance of Holstein cows milked with an automatic milking system. J. Dairy Res. 74, 40-46.

BAHR, C.; SONCK, B.; BERCKMANS, D.; VANGEYTE, J.; VAN NUFFEL, A.; VRANKEN, E.; LEROY, T.; MAERTENS, W.; SONG, X. (2008): Automatic detection of lameness in dairy cattle – Image features related to lameness. In: Proceedings of 15th Symposium and 7th Conference on Lameness in Ruminants, 289 - 293.

BENCH, J.; HOFFMANN, E.; WILSON, I. (1974): A comparison of live and videorecord viewing of infant behavior under sound stimulation. I. Neonates. Developmental Psychobiology 7, 455 – 464.

BERGSTEN, C. (2001): Effects of conformation and management system on hoof and leg diseases and lameness in dairy cows. Vet. Clin. North Am. Food Anim. Pract. 17, 1-23.

- BERRY, E.; STODDART, M.; BROUGHAN, J. (2008): Locomotion scoring of cattle using a lameness-speed index on different types of track. *Vet. Rec.* 163, 601 – 602.
- BICALHO, R.C.; CHEONG, S.H.; CRAMER, G.; GUARD, C.L. (2007): Association between a visual and an automated locomotion score in lactation Holstein cows. *J. Dairy Sci.* 90, 3294 – 3300.
- BLACKIE, N.; AMORY, J.; BLEACH, E.; SCAIFE, J. (2011): The effect of lameness on lying behaviour of zero grazed Holstein dairy cattle. *Applied Animal Behaviour Science* 134, 85 – 91.
- BLACKIE, N.; BLEACH, E.C.L.; AMORY, J.R.; SCAIFE, J.R. (2013): Association between locomotion score and kinematic measures in dairy cows with varying hoof lesion types. *J. Dairy Sci.* 96, 3564 – 3572.
- BORDERAS, T.F.; FOURNIER, A.; RUSHEN, J.; DE PASSILLÉ, A.M.B. (2008): Effect of lameness on dairy cows' visits to automatic milking systems. *Can. J. Anim. Sci.* 88, 1-8.
- BRENNINKMEYER, C.; DIPPEL, S.; MARCH, S.; BRINKMANN, J.; WINCKLER, C.; KNIERIM, U. (2007): Reliability of a subjective lameness scoring system for dairy cows. *Animal Welfare* 16, 127 – 129.
- CHAPINAL, N.; DE PASSILLÉ, A.M.; RUSHEN, J.; WAGNER, S. (2010): Automated methods for detection lameness and measuring analgesia in dairy cattle. *J. Dairy Sci.* 93, 2007 – 2013.
- COOK N.B. (2003): Prevalence of lameness among dairy cattle in Wisconsin as a function of housing type and stall surface. In: COOK N.B. (2004): The effect of housing type on lameness prevalence in Wisconsin dairy herds. University of Wisconsin-Madison School of Veterinary Medicine, Wisconsin.
- DE MOL, R.M.; ANDRÉ, G.; BLEUMER, E.J.B.; VAN DER WERFT, J.T.N.; DE HAAS, Y.; VAN REENEN, C.G. (2013): Application of day-to-day variation in behavior for the automated detection of lameness in dairy cows. *J. Dairy Sci.* 96, 3703 – 3712.
- ESPEJO, L.A.; ENDRES, M.I.; SALFER, J.A. (2006): Prevalence of lameness in high-production Holstein cows housed in freestall barns in Minnesota. *J. Dairy Sci.* 89, 3052 – 3058.
- FLOWER, F.C.; SANDERSON, D.J.; WEARY, D.M. (2005): Hoof pathologies influence kinematic measures of dairy cow gait. *J. Dairy Sci.* 88, 3166 – 3173.

- FLOWER, F.C. UND WEARY, D.M. (2006): Effect of hoof pathologies on subjective assessments of dairy cow gait. *J. Dairy Sci.* 89, 139 – 146.
- GALINDO, F. UND BROOM, D.M. (2002): Effects of lameness of dairy cows. *J Appl Anim Welf Sci.* 5, 193-201.
- GARBARINO, E.J.; HERNANDEZ, J.A.; SHEARER, J.K.; RISCO, C.A.; THATCHER, W.W. (2004): Effect of lameness on ovarian activity in postpartum Holstein cows. *J. Dairy Sci.* 87, 4123–4131.
- GONZALEZ, H., RIVEIRO, B., VAZQUEZ-FERNANDEZ, E., ARIAS, P. (2013): Metrological evaluation of Microsoft Kinect and Asus Xtion sensors, in *Measurement* 46, 1800 – 1806.
- GREEN, L.E.; HUXLEY, J.N.; BANKS, C.; GREEN, M.J. (2014): Temporal associations between low body condition, lameness and milk yield in a UK dairy herd. *Prev. Vet. Med.* 113, 63 – 71.
- GÜNTHER, M. (1991): *Klauenkrankheiten*. 5. Auflage, Gustav Fischer Verlag, Jena.
- HANSEN, M.F.; SMITH, M.L.; SMITH, L.N.; ABDUL JABBAR, K.; FORBES, D. (2018): Automated monitoring of dairy cow body condition, mobility and weight using a single 3D video capture device. *Computers in Industry* 98, 14 – 22.
- ITO, K.; VON KEYSERLINGK, M.A.G.; LEBLANC, S.J.; WEARY, D.M. (2010): Lying behavior as an indicator of lameness in dairy cows. *J. Dairy Sci.* 93, 3553 – 3560.
- JABBAR, K.A.; HANSEN, M.F.; SMITH, M.L.; SMITH, L.N. (2017): Early and non-intrusive lameness detection in dairy cows using 3-dimensional video. *Biosystems Engineering* 153, 63 – 69.
- KOFLER, J. (2014): *Diagnoseschlüssel zu Klauenerkrankungen für Klauenpfleger & Tierärzte*. Vet.med.Uni. Wien. (https://www.vetmeduni.ac.at/fileadmin/v/wiederkaeuer/Diagnoseschl%C3%BCssel_KLAUENBEFUNDE_f%C3%BCr_Klauenpfleger_Tier%C3%A4rzte_2-12-2014.pdf, abgerufen am 02.10.2017)
- MAAS J. (2009): Musculoskeletal abnormalities. In: SMITH B.P.: *Large Animal Internal Medicine*. Mosby, Inc., St. Louis, 217–231.

- MAERTENS, W.; VANGEYTE, J.; BAERT, J.; JANTUAN, A.; MERTENS, K.C.; DE CAMPENEERE, S.; PLUK, A.; OPSOMER, G.; VAN WEYENBERG, S.; VAN NUFFEL, A. (2011): Development of a real time cow gait tracking and analyzing tool to assess lameness using a pressure sensitive walkway: The GAITWAISE system. *Biosystems Engineering* 110, 29 – 39.
- MANSON, F.J. UND LEAVER, J.D. (1988). The influence of concentrate amount on locomotion and clinical lameness in dairy-cattle. *Anim. Prod.* 47,185–190.
- MÜLLER, K.E. (2011): Messen ist Wissen – Erfassung und Interpretation von Daten zur Klauengesundheit. In: FÜRLI, M. (Hrsg.): Internationale Konferenz: Prophylaxe von Herden- bzw. Produktionskrankheiten. Merkur Drucker und Kupier – Zentrum GmbH, Leipzig, 99 – 101.
- NORRING, M.; HÄGGMAN, J.; SIMOJOKI, H.; TAMMINEN, P.; WINCKLER, C.; PASTELL, M. (2014): Short communication: Lameness impairs feeding behaviour of dairy cows. *J. Dairy Sci.* 97, 4317 – 4321.
- PASTELL, M.E.; KUJALA M. (2007): A probabilistic neural network model for lameness detection. *J. Dairy Sci.* 90, 2283 – 2292.
- PASTELL, M.E.; KUJALA M.; AISLA, A.-M.; HAUTALA, M.; POIKALAINEN, V.; PRAKS, J.; VEERÄE, I.; AHOKAS, J. (2008): Detecting cow's lameness using force sensors. *Computers and Electronics in Agriculture* 64, 34 - 38.
- PETERSE, D.J. (1992): Foot Lameness. In: ANDREWS, A.H. et al.: *Bovine Medicine. Diseases and husbandry of cattle*. Oxford, 353-363.
- PLUK, A.; BAHR, C.; POUSABERI, A.; MAERTENS, W.; VAN NUFFEL, A.; BERCKMANS, D. (2012): Automatic measurements of touch and release angles of the fetlock joint for lameness detection in dairy cattle using vision techniques. *J. Dairy Sci.* 95, 1738 – 1748.
- POLITIEK, R. D.; DISTL, O.; FJELDAAS, T.; HEERES, T.; MCDANIEL, B. T.; NIELSEN, E.; PETERSE, D. J.; REURINK, A.; STRANDBERG, P. (1986): Importance of claw quality in cattle: Review and recommendations to achieve genetic improvement. Report of E.A.A.P. working group on “claw quality in cattle”. *Livestock Production Science*, Volume 15, Issue 2, Pages 133 – 152.

- POURSABERI, A.; BAHR, C.; PLUK, A.; VAN NUFFEL, A.; BERCKMANS, D. (2010): Real-time automatic lameness detection based in back posture extraction in dairy cattle: Shape analysis of cow with image processing techniques. *Computers and Electronics in Agriculture* 74, 110 - 119.
- RAJKONDAWAR, P.G.; LIU, M.; DYER, R.M.; NEERCHAL, N.K.; TASCH, U.; LEFCOURT, A.M.; EREZ, B.; VARNER, M.A. (2006): Comparison of models to identify lame cows based on gait and lesion scores, and limb movement variables. *J. Dairy Sci.* 89, 4267 - 4275.
- REHAGE, J.; JANBEN, S.; OFFINGER, J.; STRAKE, A.; MEYER, H. (2011): Lässt sich die Prävalenz von Lahmheiten in Milchviehbetrieben rasch auf unter fünf Prozent senken? In: FÜRLI, M. (Hrsg.): Internationale Konferenz: Prophylaxe von Herden- bzw. Produktionskrankheiten. Merkur Drucker und Kupier – Zentrum GmbH, Leipzig, 102 – 103.
- ROMANINI, E.C.B.; BAHR, C.; VIAZZI, S.; VAN HERTEM, T.; SCHLAGETER-TELLO, A.; HALACHMI, I. (2013). Application of image based filtering to improve the performance of an automated lameness detection system for dairy cows. In *Transactions of the ASABE. American society of agricultural and biological engineers, Kansas City, Missouri* (p. 131620675). <http://dx.doi.org/10.13031/aim.20131620675>.
- SCHLAGETER-TELLO, A.; BOKKERS, E.A.M.; GROOT KOERKAMP, P.W.G.; VAN HERTEM, T.; VIAZZI, S.; ROMANINI, C.E.B.; HALACHMI, I.; BAHR, C.; BERCKMANS, D. (2014): Manual and automatic locomotion scoring systems in dairy cows: A review. *Preventive Veterinary Medicine* 116, 12-25.
- SCHLAGETER-TELLO, A.; BOKKERS, E.A.M.; GROOT KOERKAMP, P.W.G.; VAN HERTEM, T.; VIAZZI, S.; ROMANINI, C.E.B.; HALACHMI, I.; BAHR, C.; BERCKMANS, D.; LOKHORST, K. (2015a): Comparison of locomotion scoring for dairy cows by experienced and inexperienced raters using live or video observation methods. *Animal Welfare* 24, 69 – 79.
- SCHLAGETER-TELLO, A.; BOKKERS, E.A.M.; GROOT KOERKAMP, P.W.G.; VAN HERTEM, T.; VIAZZI, S.; ROMANINI, C.E.B.; HALACHMI, I.; BAHR, C.; BERCKMANS, D.; LOKHORST, K. (2015b): Relation between observed locomotion traits and locomotion score in dairy cows. *J. Dairy Sci.* 98, 1 – 11.

- SCHLAGETER-TELLO, A.; VAN HERTEM, T.; BOKKERS, E.A.M.; VIAZZI, S.; BAHR, C.; BERCKMANS, D.; LOKHORST, K. (2018): Performance of human observers and an automatic 3-dimensional computer-vision-based locomotion scoring method to detect lameness and hoof lesions in dairy cows. *J. Dairy Sci.* 101, 1 – 14.
- SCOTT, G.B. (1989): Changes in limb loading with lameness for a number of Friesian cattle. *Br. Vet. J.* 145, 28.
- SIGH, S.S.; WARD, W.R.; LAUTENBACH, K.; MURRAY, R.D. (1993): Behaviour of lame and normal dairy cows in cubicles and in a straw yard. *Vet. Rec.* 133, 204-208.
- SOGSTAD, Å.M.; ØSTERÅS, O.; FJELDAAS, T. (2006): Bovine claw and limb disorders related to reproductive performance and production diseases. *J. Dairy Sci.* 89, 2519 – 2528.
- SONG, Y.; LEROY, T.; VRANKEN, E.; MEARTENS, W.; SONCK, B.; BERCKMANS, D. (2008): Automatic detection of lameness in dairy cattle – Vision-based trackway analysis in cow's locomotion. *Computers and Electronics in Agriculture* 64, 39 – 44.
- SPRECHER, D.J.; HOSTETLER, D.E.; KANEENE, J.B. (1997): A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology* 47: 1179 – 1187.
- STRAW, B.E., DEWE C.E., WILSON M.R. (1999): Differential diagnosis of swine diseases. In: STRAW B.E., D'ALLAIRE S., MENGELING W.L., TAYLOR D.J.: *Diseases of Swine*. Iowa State University Press, 41–86.
- TADICH, N.; FLOR, E.; GREEN, L. (2010): Association between hoof lesions and locomotion score in 1098 unsound dairy cows. *The Veterinary Journal* 184, 60 – 65.
- TASCH, U.; RAJKONDAWAR, P.G. (2004): The development of a SoftSeperatorTM for a lameness diagnostic system. *Computers and Electronics in Agriculture* 44, 239-245.
- TELEZHENKO, E. (2005): Measurements of Trackways as a Method for Assessing Locomotion in Dairy Cows. Licentiate thesis. Swedish University of Agricultural Sciences.
- TELEZHENKO, E.; BERGSTEN, C. (2005): Influence of floor type on the locomotion of dairy cows. *Animal Behaviour Science* 93, 183 – 197.
- THOMSEN, P.T.; MUNKSGAARD, L.; TOGERSEN, F.A.: (2008): Evaluation of a lameness scoring system for dairy cows. *J. Dairy Sci.* 91, 119–126.

- VAN DE GUCHT, T.; SAEYS, W.; VAN WEYENBERG, S.; LAUWERS, L.; MERTENS, K.; VANDAELE, L.; VANGEYTE, J.; VAN NUFFEL, A. (2017): Automatic cow lameness detection with a pressure mat: Effects of mat length and sensor resolution. *Computers and Electronics in Agriculture* 134, 172 – 180.
- VAN HERTEM, T.; VIAZZI, S.; STEENSELS, M.; MALTZ, E.; ANTLER, A.; ALCHANATIS, V.; SCHLAGETER-TELLO, A.A.; LOKHORST, K.; ROMANINI, E.C.B.; BAHR, C.; BERCKMANS, D.; HALACHMI, I. (2014): Automatic lameness detection based on consecutive 3D-video recordings. *Biosystems Engineering* 119, 108 – 116.
- VAN HERTEM, T.; SCHLAGETER-TELLO, A.; VIAZZI, S.; STEENSELS, M.; BAHR, C.; ROMANINI, E.C.B.; LOKHORST, K.; MALTZ, E.; HALACHMI, I.; BERCKMANS, D. (2017): Implementation of an automatic 3D vision monitor for dairy cow locomotion in a commercial farm. *Biosystems Engineering*, <http://dx.doi.org/10.1016/j.biosystemseng.2017.08.011>.
- VAN STRATEN, M.; SIANI, I.; BAR, D. (2011): Reduced test-day milk fat percentage in cows diagnosed with claw horn lesions during routine claw trimming. *J. Dairy Sci.*, Vol. 94(4), pp. 1858-1863.
- VIAZZI, S.; BAHR, C.; SCHLAGETER-TELLO, A.; VAN HERTEM, T.; ROMANINI, C.E.B.; PLUK, A.; HALACHMI, I.; LOKHORST, C.; BERCKMANS, D. (2013): Analysis of individual classification of lameness using automatic measurement of back posture in dairy cattle. *J. Dairy Sci.* 96, 257 – 266.
- VIAZZI, S.; BAHR, C.; SCHLAGETER-TELLO, A.; ROMANINI, C.E.B.; HALACHMI, I.; LOKHORST, C.; BERCKMANS, D. (2014): Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy cows. *Computers and Electronics in Agriculture* 100, 139 – 147.
- WELFARE QUALITY (2009): Assessment Protocol for Cattle. Welfare Quality Consortium, Lelystad, The Netherlands.
- WELLS, S.J.; TRENT, A.M.; MARSH, W.E.; MCGOVERN, P.G.; ROBINSON, R.A. (1993): Individual cow risk factors for clinical lameness in lactating dairy cows. *Prev. Vet. Med.* 17, 95–109.
- WHAY, H. R.; WATERMAN, A. E.; WEBSTER, A. J. F. (1997): Associations between locomotion, claw lesions and nociceptive thresholds in dairy heifers during the peripartum period. *Vet. J.* 154, 155 – 161.

WIEDENHÖFT, D. (2005): Einfluss von Lahmheiten auf die Fruchtbarkeitsleistung von Milchkühen. Dissertation, Tierärztliche Hochschule Hannover, Hannover.

WINCKLER, C.; WILLEN, S. (2001): The reliability and repeatability of a lameness scoring system for use as an indicator of welfare in dairy cattle. *Acta Agric. Scand., Sect. A: Anim. Sci.* 30, 103 – 107.

8. Anhang

Formeln zur Berechnung von Spezifität, Sensitivität, Fehlerrate und Genauigkeit

$$\text{Sensitivität} = \frac{\text{richtig positiv}}{\text{richtig positiv} + \text{falsch negativ}} \times 100$$

$$\text{Spezifität} = \frac{\text{richtig negativ}}{\text{richtig negativ} + \text{falsch positiv}} \times 100$$

$$\text{Fehlerrate} = \frac{\text{falsch positiv}}{\text{falsch positiv} + \text{richtig positiv}} \times 100$$

$$\text{Genauigkeit} = \frac{\text{richtig positiv}}{\text{richtig positiv} + \text{falsch positiv}} \times 100$$

9. Selbstständigkeitserklärung

Hiermit erkläre ich, Sarah Dieckmann, Matrikel-Nr.: 517620, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben.

Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad.

Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde.

Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin vom 5. März 2015.

Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen Promotionsbearbeiterinnen/Promotionsberatern stattgefunden hat und dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

Berlin, den 21.08.2018

Unterschrift